Wesleyan Economic Working Papers

http://repec.wesleyan.edu/
Nº: 2020-004

Age Diversity and Aggregate Productivity

Balazs Zelity

July 2020

WESLEYAN
UNIVERSITY

Department of Economics
Public Affairs Center
238 Church Street
Middletown, CT 06459-007

Tel: (860) 685-2340
Fax: (860) 685-2301
http://www.wesleyan.edu/econ

Age Diversity and Aggregate Productivity*

Balazs Zelity†

July 2020

Abstract

This research explores theoretically, empirically and quantitatively the role of age diversity in determining aggregate productivity and output. Age diversity has two conflicting effects on output. On the one hand, due to skill complementarity across different cohorts, age diversity may be beneficial. On the other hand, rapid skill-biased technological change makes age diversity costly as up-to-date education tends to be concentrated among younger cohorts. To study this trade-off, I first build an overlapping-generations (OLG) model which, in view of these two opposing forces, predicts a hump-shaped relationship between age diversity and GDP per capita. This prediction is established analytically, and also quantitatively using real-world population data in an extended computational version of the model. The prediction is then tested using country-level panel data with a novel instrument, and regional data from Europe. Moving one standard deviation closer to the optimal level of age diversity is associated with a 1.5% increase in GDP per capita. In addition, consistent with the predictions of the model, the optimal level of age diversity is lower in economies where skill-biased technological change is more prevalent.

Keywords: age diversity, education, experience, human capital, demographics, skill-biased technological change.

JEL classification: E24, O40, J24, O15.

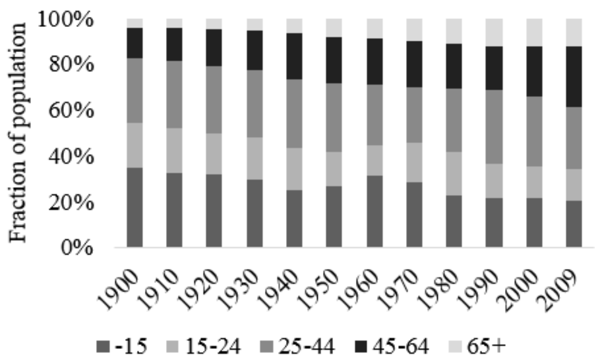
*I would like to thank Oded Galor, Stelios Michalopoulos, and David Weil for invaluable feedback and guidance. I am also grateful to Toman Barsbai, Resul Cesur, Metin Cosgel, Juan F. García-Barragán, Nils-Petter Lagerlöf, Josef Platzer, Mehdi Senouci, and participants at the RCEA Conference on Growth, Innovation and Entrepreneurship, and at Brown University seminars for their helpful comments and suggestions.

†Department of Economics, Wesleyan University, 238 Church St, Middletown, CT 06459, United States. Email: bzelity@wesleyan.edu.

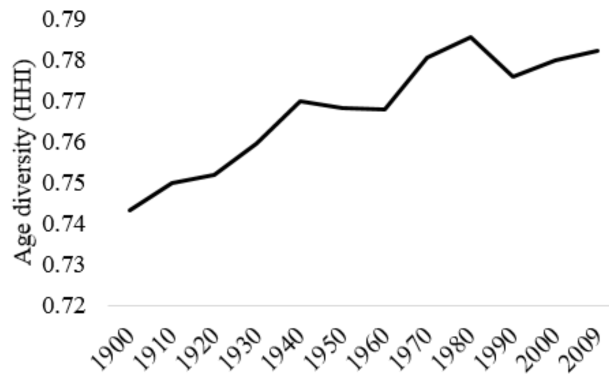
1 Introduction

There are major demographic changes ongoing in both developing and developed countries. In the developed world, most countries are experiencing population ageing, while developing countries are at various stages of the demographic transition. The effect of these demographic changes has been a fruitful area of research.[1] However, a largely ignored implication of these shifts in age structure is that they also influence a population's age diversity.

More age diverse populations are characterised by cohorts that are fairly equal in size. In less age diverse populations, some cohorts greatly outnumber others. As an illustration, Figure 1a shows how the US age structure evolved over the past century, and Figure 1b plots the corresponding age diversity.



(a) Population by age



(b) Age diversity over time

Figure 1: Recent demographic trends in the United States

This paper investigates whether changing age diversity can have important macroeconomic effects. If different age groups have different skill sets, then changing age diversity leads to changes in an economy's skill composition. For instance, if younger cohorts posses more up-to-date education, and older cohorts more experience, then increasing age diversity pushes the relative abundance of up-to-date education and experience closer to parity. This may be beneficial at first as the scarcer factor becomes more abundant, but it may be costly beyond a certain point, because the two factors need not be equally productive.

The benefit of age diversity comes from the fact that education and experience are complementary inputs into production. This means that firms need both factors in order to maximise output. Increasing age diversity ensures that both factors are available in the desired quantities. On the cost side, age diversity can be detrimental because education and experience need not be equally productive. The production technology of firms may be such that education receives a higher weight

[1] The literature is briefly reviewed later in the introduction.

in the production function. This implies that there is an optimal amount of education and experience that maximises output depending on the weight of each factor. The co-existence of a cost and a benefit to age diversity implies a hump-shaped relationship between age diversity and output per capita. Increasing age diversity is beneficial up to a point, but is detrimental beyond that.

This hypothesis is investigated in this paper in three ways. First, I build a parsimonious theoretical model to formalise the proposed relationship. Second, an extended version of the model embedded in a general equilibrium framework is simulated to quantitatively verify the prediction on real-world population data. Third, the model's implications are tested empirically. In addition, the policy implications of the findings are also explored using population projections for the 21st century.

The point of departure is an analytically tractable static OLG model with three different age cohorts. Individuals posses two types of skills: education and experience. Education is higher in younger cohorts, while experience is higher for older cohorts. One may think of younger cohorts as acquiring a more recent vintage of education. Firms use education and experience as their two factors of production. It is shown in this framework that, under some assumptions, age diversity has a hump-shaped relationship with output per capita. A key assumption behind this result is that firms do not adapt their production technologies to the prevailing demographic trends. This can happen, for instance, because rapid skill-biased technological change makes it impossible for firms to rely as heavily on experience as opposed to education as they would like.

To see if this prediction carries over to a more standard framework, the model is then extended into a full-fledged dynamic OLG model in the spirit of Auerbach and Kotlikoff (1987). The extended model features a standard household problem with endogenous investment in education, a general number of cohorts, and it is embedded in a rational expectations general equilibrium framework. It is then solved computationally, and is used to quantitatively verify the proposed hump-shaped relationship using country-level data for the 1950-2010 period.

The two main testable predictions of the theoretical set-up are as follows. First, age diversity and output per capita are expected to have a hump-shaped relationship. Second, the model also predicts that in economies where the importance of education is higher, the optimal level of age diversity is lower. This prediction suggests that skill-biased technological change is a key driver of optimal age diversity.

The next step is empirically testing these two main predictions. The empirical part of this paper is primarily based on country-level panel data for the 1950-2010 period. Further evidence is presented using European regional panel data for the 2000-2015 period. Testing the first prediction of the model amounts to regressing log GDP per capita on age diversity and squared age diversity. A hump-shaped relationship arises in accordance with the theory. Three features of the analysis ensure proper identification. First, country, year, and region (for the European data) fixed

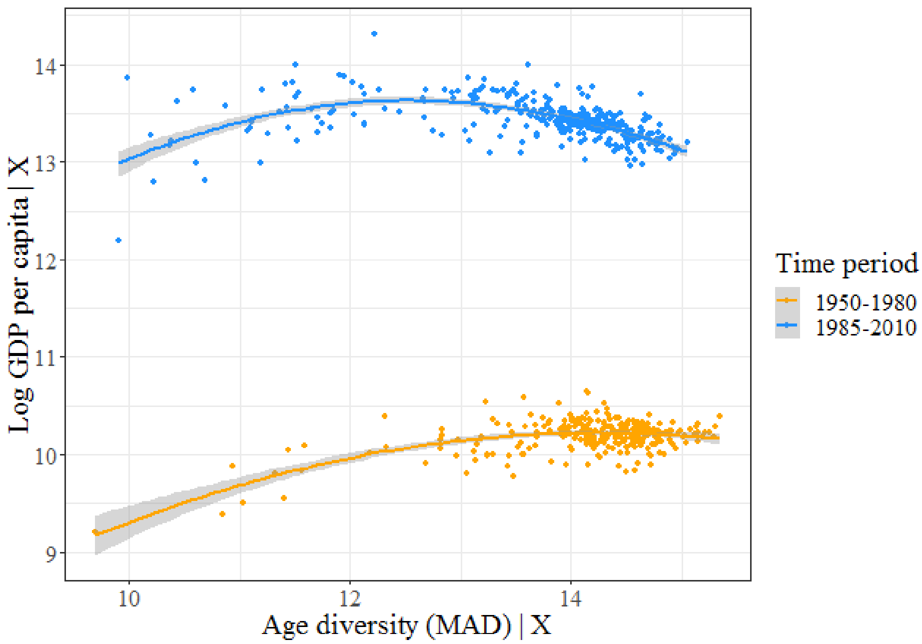


Figure 2: Conditional scatter plot of age diversity and GDP per capita

Note: This figure illustrates that age diversity and GDP per capita have a hump-shaped relationship in country-level panel data. It also shows that the optimal level of age diversity (i.e. the peak of the hump) was lower in 1985-2010 than in 1950-1980.

effects are included in all panel data specifications. This controls for all time-invariant country- and region-specific characteristics, and for all common global shocks. Second, a variety of measures of demographic structure that may be correlated with age diversity are directly controlled for. These include, for instance, mean age and dependency ratios. Third, age diversity is instrumented in the country-level analysis with past population projections from the United Nations. These past projections are purged of common shocks to age diversity and GDP per capita that could taint the results.

Figure 2 summarises the evidence for the two predictions using country-level panel data. It shows the hump-shaped association between age diversity and income per capita as well as the fact that optimal diversity was lower after 1980, a period in which returns to education were growing rapidly. The results suggest that moving from one standard deviation below (e.g. Malawi in 2010) or above (e.g. Kazakhstan in 2010) the optimal level of age diversity to the optimal level (e.g. South Africa in 2000) can increase GDP per capita by 1.5%.

To test the second prediction of the model, I perform three pieces of analysis. First, I interact age diversity and its square in the country-level analysis with measures of educational attainment. The optimal level of age diversity is found to be significantly lower in countries with higher education. Second, age diversity and its square are interacted with the growth of tertiary attainment among young people. This analysis suggests that countries with a higher growth in attainment (proxying for faster skill-biased technological change) have a lower level of optimal age diversity. Third, age diversity and its square are interacted in the European regional analysis with shares of employment in various sectors. Regions with a heavier reliance on more skill-intensive industries are found to have a lower level of optimal age diversity, once again underlying the fact that skill-biased technological change lowers the benefits of age diversity.

Finally, using the United Nations population projections by age group for 2020-2100, I project age diversity by country for the remainder of the 21st century. Using these estimates, I explore how different regions of the world will be impacted by changing age diversity in the future. In addition to this baseline scenario, age diversity can be further affected in the future by immigration policy, fertility incentives, and the retirement age, among others. A simulation suggests that all these policies can significantly affect future age diversity and, through it, income per capita, potentially reinforcing or offsetting some of the first-order effects of these policies.

Related literature

This paper contributes to three strands of literature. First, there are a set of papers that look at the economic effects of age structure. The closest study to mine in this literature is Feyrer (2007) who examines the relationship between the shares of different age groups in the workforce and aggregate productivity. Aiyar et al. (2016), meanwhile, investigate the productivity effects of population ageing. Derrien et al. (2018), Ang and Madsen (2015), and Irmen and Litina (2016) look at the interaction between ageing and innovation. Backes-Gellner and Veen (2013) is the only paper that makes age diversity its main focus, and they find evidence that age diversity is beneficial for firm productivity in creative industries in Germany. Gregory and Patuelli (2015) and Arntz and Gregory (2014) meanwhile focus on the effect of age structure on regional development within Germany. Gu and Stoyanov (2019) find that the changing age structure of advanced economies has changed what industries these countries have comparative advantage in, and has thus shifted their trade patterns.

The paper's contribution to this literature is threefold. First, the primary focus of most of the aforementioned papers has been measures such as the old- and young-age dependency ratios, or the shares of different age groups. What has so far been scarcely explored is interactions between different age groups, which is ultimately what age diversity is. The concept of age diversity is thus a new kind of demographic variable, distinct from the classic measures previously studied in the literature. Second, direct evidence about age diversity from Backes-Gellner and Veen (2013) has focused on firm-level data. This paper, in turn, focuses on macroeconomic effects. Third, I propose a formal mechanism of how age diversity and economic output are related. The literature above has for the most part focused exclusively on empirical analyses.

On a conceptual level, one of the key mechanisms driving the results in this study is imperfect substitutability between different age groups. There are a large number of papers in population economics that rely on this concept.[2] An important and oft-cited reason for imperfect substitutability in the economics literature is older workers' experience (Kremer and Thomson, 1998). Another one is cohort-specific differences in the amount and type of education obtained (Guest and Shacklock, 2005). Research in cognitive science provides more fundamental support for why different age groups may be imperfect substitutes. Karlgaard (2019) documents ample evidence that different cognitive functions peak at different stages during the life-cycle. Broadly speaking, different facets of fluid intelligence (e.g. short-term memory) peak around ages 20-30, while different facets of crystallised intelligence (e.g. vocabulary) peak around ages 40-55 (Hartshorne and Germine, 2015; Whitley et al., 2016). In addition, Gu and Stoyanov (2019) reviews an extensive literature suggesting that older people find it harder to adapt to changes and learn new skills. Firm-level evidence also suggests that different age groups are not perfectly substitutable. For example, a large number of big firms in the US technology sector have employees with a median age of around 30 (Pelisson and Hartmans, 2017).[3] Roundtree (2011) documents that age diversity, itself, was considered an imperative by 40% of respondents in a survey of 80 employers.

The paper's contribution to this literature is that an important implication of imperfect substitutability between age groups is documented both theoretically and empirically. I show that cohort-level differences in various abilities have large-scale economic consequences when aggregated.

Finally, this paper contributes to the literature on various facets of diversity and economic outcomes. Broadly speaking, this literature has found that, on the one hand, diversity may positively affect productivity by bringing together people with different perspectives and thus sparking innovation (Ashraf and Galor, 2013; Alesina et al., 2016). On the other hand, diversity can decrease productivity by fractionalising societies (Alesina et al., 2003) and making people less cooperative (Hjort, 2014). It is not surprising that these opposing forces have also been found to give rise to a hump-shaped pattern between diversity and economic growth (Ashraf and Galor, 2013).

The paper contributes to this literature in two ways. First, I consider a previously largely

[2]E.g. Freeman (1979), Welch (1979), Murphy et al. (1988), Lam (1989), Weil (1997), and Prskawetz et al. (2008) use imperfect substitutability to explain cohort-size effects on wages, while e.g. Rojas (2005), and Prskawetz et al. (2012) study the interaction between population ageing and imperfect substitutability.

[3]The young-bias of the US technology sector is also supported by survey evidence. One survey of workers in the industry showed that 43% of respondents are worried about losing their jobs due to their age (Mukherjee, 2017). Another survey found that salaries in the industry begin to fall around age 45 with people in their fifties asking for the same salaries as millennials with a mere two years of experience (Kuchler, 2017).

ignored facet of diversity. Second, age diversity is perhaps the first aspect of diversity that appears less beneficial in an innovative environment. This finding suggests that returns to diversity exhibit more heterogeneity and context-dependence than previously thought.

2 Model

In this section, a parsimonious model is developed which gives rise to the main prediction of a hump-shaped relationship between output per capita and age diversity analytically. This is an illustrative static model that attempts to showcase the logic behind the prediction. The goal here is to establish that in an economy where different age groups have different endowments of education and experience, age diversity has both costs and benefits, which gives rise to the hump-shape. A more complete and realistic model is developed in Section 3.

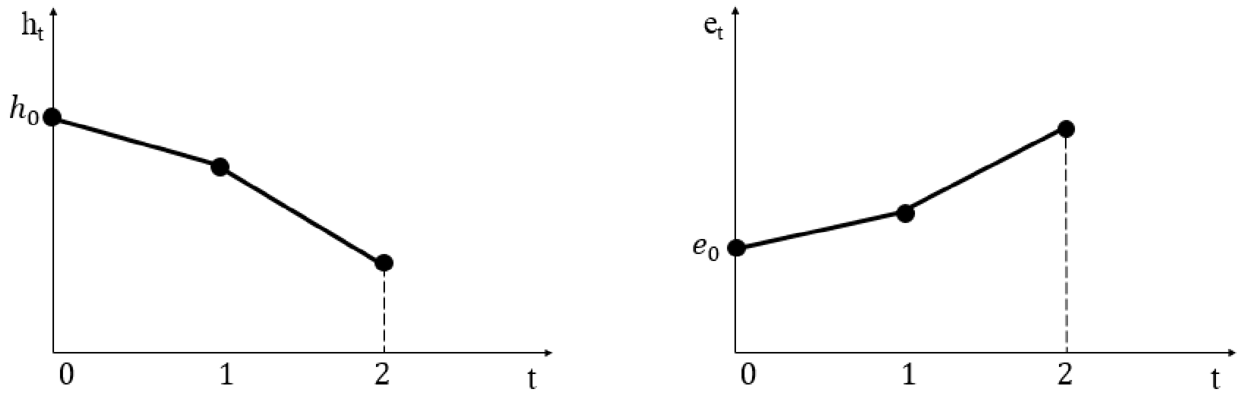
Total population is normalised to 1, and there are three age groups: 20-34-year-olds, 35-49-year-olds, and 50-64-year-olds. In this set-up, the mean (μ) and standard deviation (σ) of age fully characterise the age distribution. I treat μ and σ as exogenous parameters, and then show that output per capita has a hump-shaped relationship with the standard deviation of age, σ, which is proxying for age diversity. The first step is to express the size of the three age groups in terms of μ and σ.

Lemma 1. *The sizes of the age groups are given by*

$$n_0 = \frac{\sigma^2 + 2394 - 99\mu + \mu^2}{450}$$
$$n_1 = \frac{-\sigma^2 - 1539 + 84\mu - \mu^2}{225}$$
$$n_2 = \frac{\sigma^2 + 1134 - 69\mu + \mu^2}{450},$$

where $n_0, n_1, n_2 \in [0,1]$ *are the size of each age group from youngest to oldest.*

Output is produced using two factors of production: education and experience. Each individual is endowed with h_0 units of education and e_0 units of experience when they are born. In particular, it is assumed that education declines over the life-cycle, while experience is gained. Declining education can be thought of as younger cohorts acquiring a more recent vintage of education. In the context of the past 40 years, the higher education of younger cohorts can also be interpreted as the result of a higher initial investment in education due to increasing returns to schooling. Increasing experience can be interpreted as the accumulation of tacit knowledge, on-the-job training, and practical skills not acquired by education such as communication and management skills.



(a) Education

(b) Experience

Figure 3: The evolution of education and experience by age in the model

Figure 3 illustrates the assumption about how education and experience evolve over the life-cycle. The key assumption here is that both education and experience change less at the beginning of the life-cycle and more towards the end. So the decline in education and the gain in experience are both accelerating through the life-cycle. Mathematically, I work with the functional forms $h_t = h_0(1-t^2\delta)$ and $e_t = e_0(1+t^2\varepsilon)$, where h_t and e_t are education and experience at age t, and $\delta \in (0,1)$ and $\varepsilon \in (0,1)$ are the rates of change in education and experience.[4] These functional forms are consistent with two important empirical facts. First, they can imply a concave life-cycle earnings profile, which has been well-established in the economics literature (Lagakos et al., 2018). Second, the education depreciation curve's shape is consistent with findings about how memory declines over the life-cycle (Salthouse, 2003; Burke and Mackay, 1997; Caselli et al., 2009; Hall et al., 2009).

Aggregate education (H) and experience (E) are then given by

$$H = n_0h_0 + n_1h_1 + n_2h_2 = n_0h_0 + n_1h_0(1-\delta) + n_2h_0(1-4\delta) \tag{1}$$

$$E = n_0e_0 + n_1e_1 + n_2e_2 = n_0e_0 + n_1e_0(1+\varepsilon) + n_2e_0(1+4\varepsilon). \tag{2}$$

The aggregate production function combines education and experience using a Cobb-Douglas technology, so that total output is $Y \equiv H^\alpha E^{1-\alpha}$ for $\alpha \in (0,1)$. This also corresponds to output per capita as total population is 1.

Proposition 1 can now be derived under Assumption 1. This assumption states that the share of education in output, α, has to take an intermediate value, and the decline in education over the life-cycle cannot be too high. These conditions are necessary to ensure an interior solution for

[4]The exact functional forms are not crucial for the results. What is necessary is that the decline of education and the increase of experience are both accelerating as we move forward in the life-cycle.

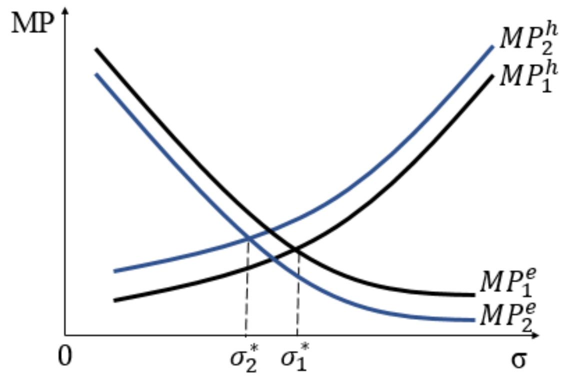


Figure 4: The marginal product of education and experience

optimal diversity.

Assumption 1. *The decline of education over the life-cycle is such that $\delta \in \left[0, \frac{1}{4}\right)$. The share of education in output is within the bounds $\frac{\varepsilon-\delta\varepsilon-X\delta\varepsilon}{\varepsilon+\delta} < \alpha < \frac{\varepsilon-X\delta\varepsilon}{\varepsilon+\delta}$, where $X \equiv \frac{1458}{450} + \frac{\mu^2}{225} - \frac{108}{450}\mu$.*

Proposition 1. *Under Assumption 1, the logarithm of output per capita has a hump-shaped relationship with age diversity as measured by σ.*

This simple model, therefore, gives rise to a hump-shaped relationship between age diversity and output per capita. The optimal level of diversity can also be calculated, and taking its derivative with respect to α yields Proposition 2.

Proposition 2. *The optimal level of age diversity as measured by σ is decreasing in the share of education in output, α.*

Proposition 2 shows that in economies with a higher reliance on education in producing output (higher α), the optimal level of age diversity will be lower. The intuition behind the two propositions is as follows. As it is apparent from Lemma 1, a mean-preserving increase in diversity (i.e. an increase in σ) redistributes people from the middle cohort (35-44) to the young (20-34) and old (45-64) cohorts. This more age-diverse population will have lower aggregate education, because $\frac{h_0+h_2}{2} < h_1$ as apparent from Figure 3a. In other words, when people are taken from the middle cohort and are split equally between the young and old cohorts, then aggregate education must decrease. Indeed, the fact that aggregate education is a decreasing function of σ is directly visible in the proof of Proposition 1. Using the same argument, it is apparent from Figure 3b that aggregate experience will increase if age diversity increases.

Decreasing aggregate education and increasing aggregate experience will be beneficial as long as the marginal product of education is lower than the marginal product of experience. However, as diversity increases, the amount of education decreases thereby increasing its marginal product, and vice versa for experience.[5] This is illustrated in Figure 4 with the MP_1^h and MP_1^e curves. At some level of diversity (σ^*), the marginal products will be equalised and if we go above σ^*, output will accordingly be decreasing in age diversity. This is the intuition behind Proposition 1. It follows that since α is the weight of education in production, it is an important determinant of where the threshold diversity lies. A higher α means education has a higher marginal product at each level of diversity and experience a lower one, which is illustrated by the MP_2^h and MP_2^e curves in Figure 4. In this case, the marginal products of education and experience will be equalised at a lower σ. This is the intuition behind Proposition 2.

The two main insights of this study have thus been established analytically in this stylised model. The key ingredients to these results are that education is concentrated among younger cohorts, while experience is concentrated among older cohorts making different age groups imperfect substitutes. In the next section, I proceed by showing that these predictions also hold in a more complete and more realistic overlapping-generations model simulated with real-world population data.

3 Simulations

I now build a more complete and realistic overlapping-generations model in the spirit of Auerbach and Kotlikoff (1987). The added complexity makes this model impossible to solve analytically, so a computational solution is necessary. The goal of this section is two-fold. First, this section shows that the main insights from the parsimonious tractable model carry over into a more realistic framework. Second, the computational model is solved with real-world population data showing that it can also match quantitative facts.

The extended model differs from the model in the previous section in three key aspects. First, the model is generalised to handle any number of cohorts, which will allow for running the model on real-world data. Second, the education investment decision is endogenised giving rise to a standard household utility maximisation problem. This verifies that the results hold even in the presence of saving, borrowing and consumption smoothing. Third, the model is embedded in a general equilibrium framework with rational expectations. With this, it is established that the results do not hinge on partial equilibrium assumptions.

The model works as follows. Households consist of several age cohorts from young to old. At the beginning of life, agents decide how much to invest in education given an exogenous return to

[5] This can also be seen mathematically in the proof of Proposition 1. The marginal product of education is $MP^h = \frac{\alpha}{h_0(1-\delta(\frac{\sigma^2}{225}+X)}$ and that of experience is $MP^e = \frac{1-\alpha}{e_0(1+\varepsilon(\frac{\sigma^2}{225}+X)}$.

education. Then as households progress through their life-cycle, their education stock (h) depreciates while their experience (e) grows. Education and experience are the two factors of production supplied by households to firms. Firms produce a single final good using these two factors, and pay households wages. Households can also save and borrow, and thus smooth their consumption over the life-cycle.

Since education and experience are complementary in the production of the final good, and since different cohorts own different quantities of these factors, there is an apparent benefit to having an age-diverse population. But it is the relative importance of education and experience in conjunction with their relative supplies provided by each cohort that determines the optimal level of age diversity. In other words, if education is more productive than experience and younger cohorts posses more of it, then it is optimal to have more younger agents than older ones. This would mean that the optimal level of age diversity is not necessarily the maximum possible level. In other words, we would expect a hump-shaped relationship between age diversity and economic output.

I now proceed to formally describe the model and present the results of the computational solution. First, the behaviour of households and firms is discussed. Then the simulation procedure is explained. Finally, the model's predictions about the relationship between age diversity and output per capita are examined.

3.1 Set-up

3.1.1 Households

Households live and work for T periods. That is, childhood and retirement are abstracted away from, and the focus is exclusively on the working-age population. Agents inelastically supply one unit of labour at a given wage rate, and they spend that wage on consumption, education investment, and saving and borrowing. Education investment is only possible in the first period of life. This set-up gives rise to the utility maximisation problem

$$\max_{h_0,\{c_t\}_{t=0}^{T},\{a_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t \frac{c_t^{1-\theta}}{1-\theta} \text{ subject to} \tag{3}$$

$$\begin{aligned} c_0 + a_1 &\leq w_0^h h_0 + w_0^e e_0 + (1+r)a_0 - \sigma h_0^2 \\ c_t + a_{t+1} &\leq w_t^h h_t + w_t^e e_t + (1+r)a_t \text{ for } t \in [1,T] \\ h_t &= (1-t^2\delta)h_0 \text{ for } t \in [1,T] \\ e_t &= (1+t^k\varepsilon)^t e_0 \text{ for } t \in [1,T] \\ a_0, e_0 &\text{ given,} \end{aligned}$$

where h_t is the education stock at time t, e_t is experience at time t, c_t is consumption at time t, a_t is saving or borrowing accumulated in period t, $r > 0$ is the interest rate paid on saving and borrowing, $\beta \in (0,1)$ is the discount factor, w_t^h and w_t^e are the wage rates per unit of education and experience, σh_0^2 is the cost of education investment with $\sigma > 0$, $\delta \in [0,1]$ is the depreciation rate of education over the life-cycle, and $\varepsilon \in [0,1]$ is the growth rate of experience over the life-cycle. The last parameter, k, controls the convexity of the experience curve, and it will be used to allow the model to match the empirically observed concave life-cycle earnings profile.

From here, it is also clear that households have two factors of production: education and experience. Their total income is a function of these two factors. Education investment is endogenously decided at the beginning of the life-cycle, but after that it exogenously depreciates. Experience is completely exogenously determined, and is growing over the life-cycle. This feature is meant to capture the idea that younger and older cohorts have different comparative advantages from the point of view of the labour market: younger cohorts possess more up-to-date and often higher education[6], whereas older cohorts have more experience.

Households perfectly observe how their education investment will affect their life-cycle wage profile. When a household makes its investment decision, it will correctly observe the future wages w_t^h and w_t^e, and it will know exactly how education depreciates and experience grows. This is a perfect-foresight equilibrium.

3.1.2 Firms

There are a large number of identical, perfectly competitive firms. They combine education and experience in a Cobb-Douglas production function. Therefore, I explicitly model education and experience as complementary inputs into the production process. On the other hand, I do not explicitly assume anything about whether different cohorts are complementary. The complementarity of cohorts will arise endogenously from the model. The profit-maximisation problem for a representative firm j is

$$\max_{h_j, e_j} h_j^{\alpha} e_j^{1-\alpha} - w^h h_j - w^e e_j \tag{4}$$

where h_j and e_j are the amount of education and experience units hired by firm j, and $\alpha \in (0,1)$ is the share of education in output. The price of the final good is normalised to 1. This is a static profit-maximisation problem meaning that in each period firms follow this behaviour. In the formulation above, the time subscripts are omitted, but in any given period t, w^h and w^e refer to

[6] In reality, the education of younger cohorts can be higher for two reasons. First, they receive a more recent vintage of education, which is more productive and relevant. Second, in the presence of skill-biased technological change, they are likely to invest more in education than older cohorts did.

that period's wage rates, i.e. to w_t^h and w_t^e.

It follows that aggregate output in the economy is given by $Y_t = H_t^{\alpha_t} E_t^{1-\alpha_t}$, where H_t and E_t are the aggregate education and experience demanded by firms, respectively.

3.1.3 Equilibrium

The equilibrium of the model satisfies the following criteria.

- Households maximise utility according to (3).
- Firms maximise profits according (4).
- The wage rates w_t^h and w_t^e are determined so that the labour markets for education and experience clear. The output market clears by Walras' law.

There are no general results as to whether such an equilibrium exists and is unique in these types of OLG models (Laitner (1990)). Existence and uniqueness are usually visually verified by plotting supply and demand functions. Figure C.1 illustrates that supply and demand functions in the model are well-behaved. This shows that an equilibrium exists at the parameter values I consider, and that it is likely unique.

3.2 Simulation procedure

I now describe the solution of the model, and how I evaluate it on real-world population data. The goal is to run the model on all the country-level population structures that have been observed in the world between 1950-2010. Then I plot each population structure's age diversity against its model-predicted GDP per capita, and see if a hump-shape arises. To do this, I implement the following simulation procedure.

1. Pick a set of parameters as summarised in Table D.1.
2. Get the size of each age group, $\{N_k\}_{k=0}^T$, from country-level panel data for a large number of country-year pairs.
3. Solve the model for each country-year pair given the parameters and population structures $\{N_k\}_{k=0}^T$.
4. Calculate output per capita in the model for each empirically-observed population structure $\{N_k\}_{k=0}^T$. I.e. output per capita is calculated for the empirically-observed age structure of every single country from 1950 to 2010. There will thus be an output per capita prediction for many different empirically-observed age structures.

5. Plot a measure summarising the age diversity of each country-year observation against model-predicted output per capita to examine the relationship between the two variables.

As an example of this process, consider the following. First, parameter values are picked. Then all that is needed to be calculate output per capita are the N_k, which refer to the size of group k in the working-age population. The working-age population is divided into nine five-year age groups: from 20-24 year-olds to 60-64 year-olds. This is the finest resolution data that is widely available at the country level. In principle, output per capita could be calculated at any set of N_k: the case in which all groups have equal share, or the case in which the entire population is 20-24 years old, etc. In practice, the focus is instead on the empirically-observed shares of groups.

For instance, I observe the share of each age group in Germany in 1960, so I will calculate output per capita for that N_k combination. I also observe this for Germany in 1965, so I will calculate output per capita again for that potentially different N_k combination. I do this for the N_k combination of all countries for the 1950-2010 period. I then have a model-predicted GDP per capita for each N_k combination that has been observed in the world since 1950.

The last step is to summarise each N_k combination as one number measuring the age diversity of that population. For instance, if Germany in 1960 consists only of 20-24 year-olds, then it is not very age-diverse. But if in 1965 each age group has an equal share, then Germany is very age-diverse. Measures of age diversity are discussed in detail in Section 4.2. For now, it is sufficient to know that the measure used in the model simulations is the mean absolute difference (MAD) of the population's age structure. Intuitively, this metric measures the expected age difference between two randomly picked people from the working-age population. In the 1950-2010 country-level panel data, its value ranges from 9.694 to 15.405.

3.2.1 Model solution

The only remaining question is how to solve the model itself in Step 3 of the simulation procedure. A key difficulty comes from the fact that the population structure $\{N_k\}_{k=0}^{T}$ in any given simulation need not be consistent with a demographic steady state.[7] Therefore, the model needs to be solved out of steady state.

To achieve this, I consider a transition in which the population structure departs from its steady state, gradually turn into the empirically-observed population structure, and then gradually change back to the steady state.

As an example, consider a simplification of the model with only two age groups: young and old. Then in the initial steady state, we would have 50 young people and 50 old people. In the final

[7] For instance, if there are are five 20-24-year-olds, four 25-29-year-olds, and five 30-34-year-olds, then the demographics are not in a steady state. A constant population growth or decline would be a steady state, but a population in which the relative size of age groups is jumping around is not.

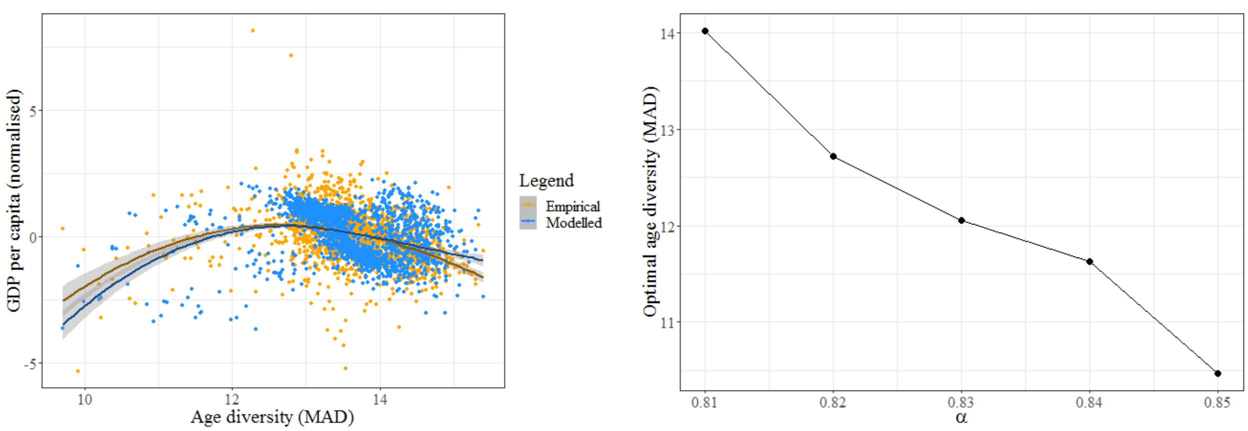
Table 1: Illustration of out-of-steady-state dynamics

	Initial SS	Transitory period 1	2	3	Final SS
Young people	50	70	30	50	50
Old people	50	50	70	30	50

steady state, we would have 50 young people and 50 old people as well. Suppose we want to see the model's predicted output per capita for a population structure of 30 young people and 70 old people. Then the evolution of the population structure would follow the path laid out in Table 1. In this case, we would be interested in output per capita in the model in the second transitory period, as this is when the population structure is the empirically-observed one. When I implement this algorithm, there are more than just two age groups as the population structure data is in five-year age groups. This means that the transitory period will be longer, and the transition more gradual. But otherwise what I do is identical to the example in Table 1.

I solve the household and firm problems as well as find the market-clearing wages using common procedures such as value function iteration, and guessing an initial market-clearing wage and converging to the actual one (Auerbach and Kotlikoff (1987)). The details of these solution methods are summarised in Appendix B.

3.3 Simulation results



(a) Modelled and empirical relationship

(b) Sensitivity of optimal diversity to α

Figure 5: Results of simulations

The model produces a hump-shaped relationship between age diversity and GDP per capita in the simulations as Figure 5a shows. It is also apparent from the figure that the modelled relationship can match the empirical relationship very closely. The empirical relationship is explored in much more detail in Section 5. This model prediction, therefore, shows that Proposition 1 carries over to the more realistic model under reasonable parameter choices. It also shows that the model is capable of producing a hump similar to the one empirically observed when estimated on real-world population data. This simulation result in conjunction with Proposition 1 gives rise to the first testable prediction.

Testable Prediction 1. *The relationship between age diversity and GDP per capita is hump-shaped when keeping country-specific characteristics constant.*

It is important to note that both Proposition 1 and the simulations keep all parameters constant across countries, so the relationship predicted by the models is purely driven by age diversity. This is why Testable Prediction 1 clarifies that country-specific characteristics need to be held constant.

To conclude this discussion of the simulations, let us consider whether the result from Proposition 2 carries over to the simulations as well. To do this, I examine the sensitivity of the optimal level of diversity to the importance of education in the model as measured by α. The results are shown in Figure 5b. The higher the productivity of education is (α), the lower the optimal level of age diversity is. This indicates that skill-biased technological change lowers the optimal level of age diversity. The intuition is that skill-biased technological change increases the relative importance of education, but since younger agents are better endowed with education, the optimal level of diversity decreases. In combination with Proposition 2, this result gives rise to the second testable prediction.

Testable Prediction 2. *The optimal level of age diversity is lower in countries with higher α (returns to education), keeping all else constant.*

It is also worth noting that the model's calibration implies a concave life-cycle earnings profile in accordance with the literature (Lagakos et al., 2018). This is illustrated in Figure C.3. This completes the description of the model and its predictions. In the upcoming two sections, I proceed with the empirical analysis of the two testable predictions.

4 Data and empirical strategy

The goal now is to examine whether the two testable hypotheses are backed by the data. In this section, the empirical strategy is discussed first, then the measurement of age diversity is explained.

4.1 Empirical strategy

To examine the validity of the hypotheses, I estimate panel-data regressions of the form

$$y_{ct} = \alpha_0 d_{ct} + \alpha_1 d_{ct}^2 + X_{ct}\beta + \gamma_c + \delta_t + \varepsilon_{ct}, \tag{5}$$

where y_{ct} is log GDP per capita in country c in year t, d_{ct} is the age diversity of the working-age population (20-64-year-olds), X_{ct} is a matrix of control variables, and γ_c and δ_t are country and year fixed effects. The residuals are denoted ε_{ct}, and they are clustered at the country-level to account for autocorrelation in the time series within countries.

Testable Prediction 1 says that $\alpha_0 > 0$ and $\alpha_1 < 0$. It is easy to ascertain whether this is the case after estimating (5). Testable Prediction 2 says that the optimal level of age diversity, given by $d^* \equiv -\frac{\alpha_0}{2\alpha_1}$, is lower when returns to education are higher. This prediction is examined by interacting d_{ct} and d_{ct}^2 with measures proxying for returns to education.

4.1.1 Data

Equation (5) is estimated using both country-level and European regional (NUTS2) panel data. In the latter case, I also add regional fixed effects and cluster the standard errors at the regional level. The country-level data has observations from 1960 to 2010 in five-year intervals. The European data is annual and spans 2000-2015.

The main dependent variable is GDP per capita, which for the country-level data is obtained from the Penn World Table (Feenstra et al. (2015)), and for Europe from Eurostat. The control variables are mean age, the young- and old-age dependency ratios, and the share of 20-29, 30-39, 40-49, 50-59 and 60-64 year-olds in the working-age population. The source for the demographic variables on the country-level is the United Nations (2017), on the European level it is the Eurostat. To examine Testable Prediction 2, two proxies are used for returns to education (α): (1) the fraction of people with more than primary educational attainment, and (2) the growth rate of the fraction of 20-24-year-olds with tertiary education relative to five years ago to proxy for growth in returns to education. The source of data is Lee and Lee (2016). The measurement of age diversity is discussed in Section 4.2. A more detailed description of all variables can be found in Appendix E.

4.1.2 Identification

In the country-level regressions, reverse causality is unlikely to be an issue for specification (5), because current GDP per capita cannot easily affect current age diversity without considerable cross-country migration. It may, however, be an issue in the regional analysis. Another concern may be omitted variable bias. It could be for instance that past GDP per capita affected fertility patterns and thus current age diversity, and it also affected current GDP per capita. This would, however, be a time-invariant omitted variable bias, and all such biases are effectively controlled for by the comprehensive country and region fixed effects. These fixed effects also ensure that factors such as institutions and culture are unlikely to be driving the results. In addition, the year fixed effects ensure that any common global shocks to GDP in a given year are controlled for as well.

The biggest threat to identification are, therefore, country-specific (or at least not globally common) time-varying factors that are correlated with both GDP per capita and age diversity. Perhaps the most obvious examples here are other measures of demographic structure such as the old- and young-age dependency ratios, mean age, or the share of various age groups in the population. Controlling for such measures can help resolve most of these issues, but it still cannot necessarily be ensured that all alternative explanations have been excluded. In order to overcome this issue, age diversity is instrumented using past population projections from the United Nations.

The United Nations has been preparing population projections for decades. The identification strategy involves calculating age diversity from old UN population forecasts instead of actual population numbers. This gives a measure of "predicted age diversity" that is purged of unexpected shocks, and is thus conceivably uncorrelated with the error term in the regression.

The oldest UN projections to have five-year age groups up to the present day are, to my knowledge, from 1982 (Keilman (1998), United Nations (1985)). With this, predicted age diversity can be calculated from 1985 to 2015 in five-year intervals. Replacing actual age diversity in (5) with predicted age diversity (a reduced-form estimate) and using predicted age diversity as an instrument (an IV estimate) are the two ways in which I attempt to identify the causal effect of age diversity on GDP per capita.

Finally, to deal with the potential unit-root problem in (5), the models are estimated both in levels and first-differences.

4.2 Measuring age diversity

This section has three goals: (1) to explain how age diversity is measured, (2) to show that age diversity is a novel dimension of population structure, and (3) to illustrate how age diversity can act as a summary measure of the population age structure.

There are a number of ways to measure age diversity. Ultimately, all rely on demographic data, so the source for calculating age diversity is the same as for other demographic variables. Table D.2 summarises the six metrics of age diversity that are considered in this paper. These are the mean absolute difference (MAD), the standard deviation (SD), the Gini coefficient, the Herfindahl-Hirschman Index (HHI), the Generalised Entropy Index (GE), and the Atkinson Index. These measures encompass a very wide spectrum of conceivable metrics for age diversity. My preferred measure is the mean absolute difference (MAD) because of its intuitive interpretation: it measures the expected age difference between two randomly picked people from the working-age population. But I examine robustness to all measures.

To get a feel of how these age diversity metrics are distributed, Figure C.4 shows the histograms of all six measures, while Figure C.5 maps age diversity at both the country-level and for European regions. In addition, Figure C.6 shows that if the European regional measures of age diversity are aggregated to the country-level, then they are largely consistent with the country-level numbers from the UN. This confirms that the two data sources are consistent with each other, and thus reasonably reliable. Further, Table D.3 shows that the six different measures of age diversity are largely positively correlated with each other. The two exceptions are the HHI and to a lesser extent the Gini. The HHI is a special case, because it is an index meant for categorical data: it doesn't take the distance between any two age groups into account. In other words, it assumes that a 20-year-old is as different from a 25-year-old as from a 40-year-old. The other indices do not make this assumption.

Table 2: Within-country correlation of age diversity with demographics

	Old AD	Young AD	Mean age	20-29	30-39	40-49	50-59	60-64
MAD	0.286	0.127	0.487	0.141	0.678	0.004	0.657	0.653
SD	0.246	0.205	0.341	0.027	0.682	0.170	0.558	0.616
Gini	0.216	0.504	0.393	0.673	0.492	0.559	0.078	0.033
HHI	0.356	0.298	0.871	0.828	0.164	0.599	0.720	0.561
GE	0.053	0.348	0.069	0.299	0.750	0.229	0.347	0.357
Atkinson	0.008	0.385	0.018	0.379	0.722	0.298	0.276	0.304

Note: "Old" and "Young AD" refer to the old- and young-age dependency ratios. "Mean age" is the mean age of the working-age (20-64) population. The remaining columns refer to the share of various age groups (e.g. 20-29-year-olds) in the working-age population. The GE index is calculated for $\alpha = -1.4$, the Atkinson index for $\varepsilon = 2$.

While the different measures of age diversity are largely correlated with each other, Table 2 shows that age diversity as a whole is not strongly correlated with common demographic characteristics. As seen in the table, none of the measures of age diversity, with the exception of HHI, have a very strong within-country correlation with demographic indicators commonly considered in other studies. Age diversity therefore measures a previously unexplored dimension of demographics.

Tables D.4-D.5 illustrate what aspects of population structure the measures of age diversity can pick up. Table D.4 shows how population structure and age diversity varied in the United States form 1970 to 2015. According to most measures, age diversity initially increased as the number of older people in the working-age population went up. Then as the baby boomers started entering working age around 1980, age diversity went down as this group started to dominate the workforce. From 2000 onwards, age diversity has once again been on the rise as most baby boomers have not retired yet, and continued population growth created a mostly flat age distribution for 20-64-year-olds.

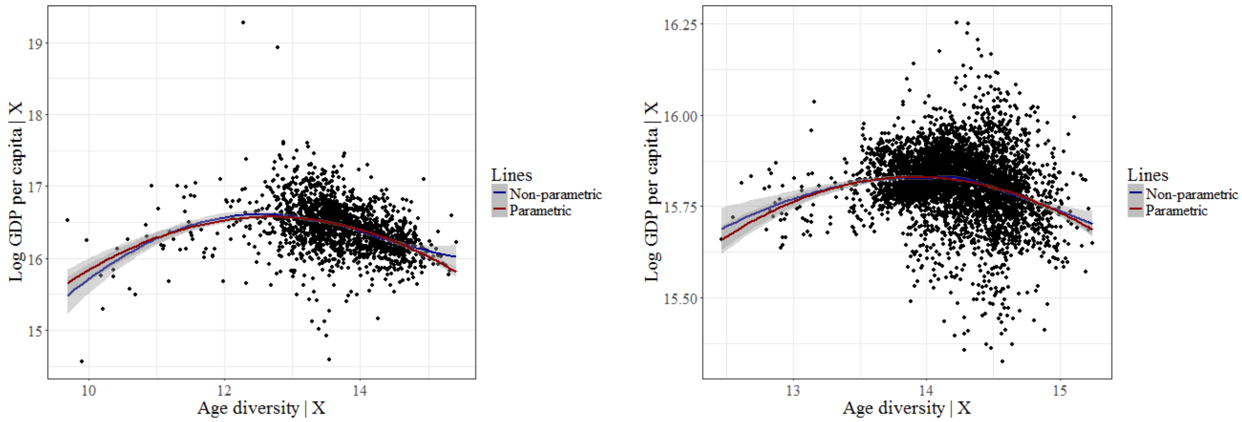
Table D.5 illustrates the connection between population structure and age diversity in an international context. The United Arab Emirates in 1980 ("ARE") had the lowest age diversity in the entire 1950-2015 sample. This is due to a single-peaked age distribution in which most people are concentrated around a single age. The Solomon Islands in 1985 ("SLB") represent a smoother downward-sloping distribution common in developing countries. The diversity of such a population is higher than the single-peaked distribution but in general still rather low – it is around the 25th percentile of the data. The Bahamas in 2005 ("BHS") exhibit an age distribution that is partially flat and partially downward-sloping. This is common in countries that had recently gone through the demographic transition, and is associated with roughly the median level of age diversity. France in 2010 ("FRA") and Lithuania in 2015 ("LTU") represent levels of diversity from the upper end of the distribution. France is largely flat, while Lithuania exhibits a double-peak, which gives rise to an even higher level of diversity.

5 Empirical results

This section presents the results of the empirical analysis. It begins by a description of the results on Testable Prediction 1, that is whether there is indeed a hump-shaped relationship between age diversity and income per capita. Then, this section discusses Testable Prediction 2 on whether proxies for skill-biased technical change affect the optimal level of age diversity.

5.1 Age diversity and income per capita

Table D.6 shows the estimation of Equation (5) using country-level panel data from 1950 to 2010. Age diversity has a clear hump-shaped association with GDP per capita in all columns. This means the relationship is robust to the gradual inclusion of demographic controls such as the old- and young-age dependency ratios, mean age, and shares of various age groups in the population. The coefficient estimates are fairly stable across specifications suggesting little to no selection on unobservables bias. The peak of the hump varies between 12.512 and 13.809 depending on the specification. Figure 6a shows a conditional scatter plot corresponding to Column (5) of Table D.6. The hump-shaped relationship is apparent from both the parametric and the non-parametric curves.



(a) Countries, 1950-2010

(b) European regions, 2000-2015

Figure 6: Conditional scatter plot between age diversity and GDP per capita

The estimates in Column (5) suggest that moving from one standard deviation (0.695) above or below the optimal level of age diversity (12.681) to the optimal level is associated with a 5.1% higher income per capita. This estimate, however, suffers from a few issues. First, due to the time series dimension of the data, it might have a unit root. To alleviate this concern, the specification is estimated in first differences in Table D.7. For the remainder of this paper, all tables are presented in first differences, but the results also hold in levels. The second issue with the estimates in Table D.6 is potential endogeneity, which is addressed later.

Table D.7 therefore presents the first-differenced estimation of (5). The conclusions are broadly the same as for Table D.6: a robust and significant hump-shaped association arises. The estimated peak of the hump now shifts upward to the range 13.230-14.413. The preferred estimate from Column (5) is 13.286, which is included in the range of peaks estimated in Table D.6 as well. These coefficients suggest that moving from one standard deviation above or below the optimal level of age diversity to the optimal level is associated with a 2.3% higher GDP per capita.

Table D.8 shows the first-differenced specification using predicted age diversity and predicted demographic controls instead of actual ones. The robust and significant hump-shape remains despite the much smaller sample size. Given the reliance on UN population projections from 1982, predicted diversity data is only available from 1985 onward, necessarily restricting the time span of the sample to 1985-2010. The estimated peak in the preferred specification of Column (5) is 13.313, which is close to the peaks estimated in the other specifications. Moving from one standard deviation above or below the optimal level of diversity to the optimal level is now associated with a 1.5% higher GDP per capita. This is the most complete and preferred specification evaluating Testable Prediction 1.

Finally, Table D.9 presents the first-differenced results when instrumenting actual age diversity with predicted age diversity. In this table the sample size necessarily plummets further as only those observations remain for which there are observations of both actual and predicted age diversity. Nevertheless, the robust hump-shape prevails and remains significant in most columns.

Table D.10 shows the population structure of the five countries closest to the estimated optimal level of diversity. The countries are the Comoros in 1990, South Africa in 2000, Costa Rica in 1950, Ethiopia in 1970, and Libya in 1990. Their population structures are very similar: they all exhibit moderately downward-sloping age structures. In other words, these are all countries where younger workers outnumber older workers to a certain, but not extreme, degree. As evident from Table D.4, the population structure of the United States most closely resembled the optimal one in 1995. This also happens to coincide with a period of rapid growth in the US.

Decomposing the effect

In accordance with the theory, Tables D.11-D.12 show that the main channel driving the hump-shaped relationship between GDP per capita and age diversity is productivity. These tables regress total factor productivity (TFP), capital per capita, and the human capital index from the Penn World Table on age diversity and its square. These three dependent variables are the major components of GDP per capita. Therefore, this exercise can be thought as a decomposition of the effect of age diversity on income per capita. A significant hump-shape is only obtained for TFP suggesting that productivity is the key channel. Table D.11 is a first-differenced specification using actual age diversity, essentially corresponding to Table D.7. The odd columns of Table D.11 estimate the same unconditional model as Column (1) of Table D.7. The even columns estimate the fully-controlled specification from Column (5) of Table D.7. It is interesting to note that the optimal level of age diversity for maximizing TFP is 13.300 in Column (2) of Table D.11, which is very close to the optimal level for GDP per capita at 13.313.

Table D.12 meanwhile runs the same specifications but using predicted age diversity instead of the actual one. This table, therefore, corresponds to Table D.8. The sample size now plummets, but a significant hump-shaped relationship is once again only obtained for TFP, and not for capital per capita or human capital. The same results hold when actual age diversity is instrumented by predicted age diversity as seen in Table D.13.

Alternative measures of age diversity

As mentioned previously, age diversity can be measured in a variety of ways. In particular, I have identified six measures. Tables D.14-D.17 show that the hump-shaped relationship from Testable Prediction 1 is robust to a number of measures of age diversity. Tables D.14-D.15 show the estimates with actual diversity, Tables D.16-D.17 with predicted diversity, and Tables D.18-D.19

with instrumented age diversity. The conclusions from all tables are largely the same. Measures such as standard deviation, the Gini coefficient, and the Generalised Entropy and Atkinson indices all give rise to the same hump-shaped relationship as the mean absolute difference. For these indicators the optimal level of diversity is in all but one case somewhat below their median level – which is the case for MAD as well. The hump is therefore not sensitive to what measure of age diversity is used.

In order to calculate the Generalised Entropy and Atkinson indices, one has to pick a value for the parameters α and ε, respectively. As Tables D.15, D.17 and D.19 show, several values have been tried. In addition, I searched for the values of α and ε that minimise the sum of squared residuals. The results of this search are summarised in Figure C.7, and show that an α of -1.4 and an ε of 3.0 provide the best fit. The figures also show that the significant hump-shape is not overly sensitive to the parameter values.

The last thing to note about the alternative measures is that the hump-shape is not present in a robust way for the Herfindahl-Hirschman Index. This is not entirely surprising as this measure effectively discards a large amount of data that the other indices take into account.[8] In particular, as discussed in Section 4, the HHI is a categorical measure that doesn't take the age differences between any two groups into account. Therefore, it can be expected that it will miss an important component of age diversity. In Tables D.14, D.16 and D.18, three different measures of the HHI were calculated to investigate whether the relationship is perhaps sensitive to how an age group is defined. The three measures divide the population into five-, ten-, and fifteen-year age groups. The hump-shape is only there consistently for the five-year groups, perhaps because there is little variation across time and countries when ten- and fifteen-year age groups are considered.

European regional data

To further show the robustness of the results, Equation (5) is estimated using regional (NUTS2) data from Europe. The specification now spans the period 2000-2015, and includes region fixed effects in addition to the country and year fixed effects. Table D.20 shows the results of this regression including all the demographic controls for the six different measures of age diversity. A hump-shape arises in this different data set as well regardless of the measure of age diversity used. The peak for MAD is 13.913, which is a bit higher than the peak estimated on the country-level data, but is in the same ballpark. I explore what determines the position of the peak in more detail in Section 5.2. The results for the Generalised Entropy and Atkinson indices are once again insensitive to the values of α or ε as shown in Figure C.8. The size of the effect is similar to the one estimated in the country-level data: moving one standard deviation (0.374) above or below the

[8]This is also visible in the formula for HHI in Table D.2: the calculation of HHI does not depend on the age of a given group, a_i.

optimal diversity to the optimal level is associated with a 1.2% higher GDP per capita. Figure 6b shows the conditional scatter plot between MAD and GDP per capita.

In addition to estimating Equation (5) on NUTS2 regions, I also aggregated the regional data to the country-level, and ran the specifications with country-level variables. The results are summarised in Table D.21, and they broadly reinforce the findings so far: a significant hump-shape arises, which is robust to demographic controls and to the choice of the age diversity measure. Just as for the country-level analysis, the only measure of age diversity that is not clearly associated with GDP per capita in a hump-shaped fashion is the categorical Herfindahl-Hirschman Index.

Naturally, the analysis involving European regions is somewhat correlational[9] as "predicted age diversity" is unavailable in European regional data. One issue might be that rural-urban differences in age diversity are driving the results. To some extent, this hypothesis can be tested. For a given country, the Generalised Entropy index can be decomposed into two terms: (1) between-region age diversity, and (2) within-region age diversity. Ultimately, the size of each of these components tells us where age diversity comes from in a given country. Does it come from the fact that regions have very different age structures (between-region), or from the fact that within each region the age structure is diverse (within-region)? A rural-urban divide would be more consistent with between-region diversity. Figure C.9 shows the percentage of age diversity (as measured by GE) that is contributed by within-region diversity in each country in 2015. It is apparent that the vast majority of diversity (more than 98% for all countries) comes from within-region differences suggesting that a rural-urban divide is not the main driving force behind age diversity.

Robustness to mean age

An important implication of Proposition 1 is that the hump-shaped relationship between age diversity and GDP per capita should be observable while holding the mean age constant. While all of the previous analysis included specifications that control linearly for the mean age of the working age population, Tables D.22-D.23 use more flexible controls for this variable. In particular, all columns in these two tables control for up to a fifth-degree polynomial of mean age. Table D.22 implements this for the country-level analysis, and Table D.23 for the European regional analysis. The results are qualitatively and quantitatively unchanged by this robustness check.

5.2 Determinants of the hump

After having established a robust hump-shaped pattern, I now turn to examining what determines the position of the optimal level of age diversity. The model has already shed some light on this

[9] While I cannot address all endogeneity concerns, I do have a comprehensive set of year, region, and country fixed effects as well as demographic controls.

issue suggesting that returns to education is an important determinant of the peak of the hump. This result was summarised above in Testable Hypothesis 2. In this section, this hypothesis is tested in both country-level and EU data.

5.2.1 The effect of returns to education

Testable Hypothesis 2 states that higher returns to education (α) lower the optimal level of age diversity. The aim is now to investigate this hypothesis in the country-level data. To start, returns to education needs to be quantified. Unfortunately, there is no panel data on returns to education for the time period the rest of the data spans.[10]

To overcome this issue, returns to education are proxied by two measures. First, I consider the fraction of people with more than primary educational attainment constructed from the data set of Lee and Lee (2016). It is to be expected that as returns to education increase, so does attainment. Hence there should be a reasonable correlation between returns to education and this measure. Second, I also look at the growth rate of educational attainment by calculating the ratio

$$\frac{\text{\% 20-24-year-olds enrolled in tertiary education in year } t}{\text{\% 20-24-year-olds enrolled in tertiary education in year } t-5}.$$

This captures how much more education younger people chose to acquire relative to the cohort before them, and as such it can be seen as a good proxy for changes in returns to education. I focus on 20-24-year-olds, because the cohort size in the model is also five years. The results do not change if the above ratio is calculated instead for 20-29-year-olds or 20-34-year-olds. To obtain this measure, age-specific levels of enrollment are calculated following Ang and Madsen (2015). The ultimate source for this data is also Lee and Lee (2016).[11]

To test whether the optimal level of age diversity changes in response to these two variables, specification (5) is estimated with the addition of an interaction between age diversity (d_{ct}) and the measures of returns to education (r_{ct}),

$$y_{ct} = \alpha_0^0 d_{ct} + \alpha_0^1 d_{ct} r_{ct} + \alpha_1^0 d_{ct}^2 + \alpha_1^1 d_{ct}^2 r_{ct} + \xi r_{ct} + X_{ct}\beta + \gamma_c + \delta_t + \varepsilon_{ct}. \tag{6}$$

This specification therefore interacts r_{ct} with both the linear and the quadratic d_{ct} terms to allow

[10] The most complete data, to my knowledge, is from Psacharopoulos and Patrinos (2018), but it unfortunately does not go back far enough in time.

[11] While the data set of Cohen and Leker (2014) could serve as an alternative, it only reports average years of schooling, and is only available at 10-year intervals.

for more generality. The optimal level of diversity in this specification is given by

$$d_{ct}^{*} = -\frac{1}{2}\frac{\alpha_0^0 + \alpha_0^1 r_{ct}}{\alpha_1^0 + \alpha_1^1 r_{ct}}.$$

Table D.24 summarises the results from estimating specification (6). The first three columns show the estimates when interacting age diversity with the share of people with more than primary educational attainment. The last three columns show the estimates when interacting with the growth rate in returns. Columns (1) and (4) show estimates with actual age diversity, Columns (2) and (5) use predicted age diversity, and Columns (3) and (6) instrument for age diversity using predicted age diversity. Column (4) indicates that the growth in returns to education is a significant determinant of the optimal level of age diversity. Column (1) shows that returns to education may also affect optimal diversity but the result is not significant. Unfortunately, due to low sample size issues the predicted and instrumented columns are not accurately estimated either.

To show how the coefficients and optimal levels of diversity vary as returns to education change, Table D.25 presents the linear and quadratic coefficients on age diversity as well as the implied optimal diversity for a wide range of quantiles of returns to education. For instance, the column corresponding to the 20th quantile shows the coefficients and the peak assuming that r_{ct} is equal to its 20th quantile. The numbers in Table D.25 are in broad agreement with Testable Prediction 2. First of all, the hump-shape is robust across the entire range of r_{ct}: the linear coefficients are positive, the quadratic ones are negative. Furthermore, the optimal level of diversity is decreasing as returns and growth to returns increase. This is in accordance with Testable Hypothesis 2.

5.2.2 Optimal age diversity by sector

As a further check of Testable Hypothesis 2, I turn to the regional data from the EU and estimate (6) while interacting age diversity with the share of employment in four economic sectors. This specification tells us whether regions that are more reliant on certain sectors have a higher or lower level of optimal age diversity.

To link this with Testable Hypothesis 2, I focus on four sectors whose skill-intensity can be reasonably well ranked. From most to least skill-intensive these are: high-technology manufacturing and knowledge-intensive services, all manufacturing, low-technology manufacturing, and agriculture. In accordance with Testable Hypothesis 2, it is expected that a higher reliance on more skill-intensive sectors is associated with lower levels of optimal age diversity. This is because returns to education can be expected to be higher in regions with a higher reliance on skill-intensive industries.

Table D.26 presents the results of this specification. In all columns the hump-shaped relationship between age diversity and GDP per capita is retained. Column (1) shows that the optimal age diversity decreases as a region's reliance on high-technology industries increases. Columns (2)-(3) show that the optimal diversity also decreases as reliance on manufacturing and low-technology manufacturing increases, but to a much smaller extent than in the case of high-technology industries. Finally, Column (4) shows that in the least skill-intensive industry, agriculture, optimal diversity actually increases as employment in this industry grows. In addition, Table D.25 shows how the linear and quadratic coefficients on age diversity as well as the optimal level of age diversity vary as employment in a given sector increases gradually from its sample minimum to its sample maximum. These results are broadly consistent with Testable Hypothesis 2.

6 Policy implications

This section discusses the policy implications of the findings from Section 5. Given the specification in (5), it is clear that holding everything else constant, the effect of age diversity in any year $t > 2015$ on income per capita relative to 2015 is

$$\frac{GDP_t}{GDP_{2015}} = e^{\alpha_0(d_t - d_{2015}) + \alpha_1(d_t^2 - d_{2015}^2)}.$$

Using this relationship, I simulate the economic effect of policies that impact age diversity. It is important to note that the simulations only document the economic effect of the policy through the channel of age diversity as we are holding everything else constant. Other channels may reinforce or offset these effects.

Four experiments are conducted. First, using population projections by age group for the period 2015-2100 from the United Nations (2017), projected age diversity is calculated by country for the remainder of the 21st century. I then predict the effect these "natural" changes in age diversity will have on GDP per capita. Second, I simulate the effect of a continuous immigration policy allowing 20-24-year-old immigrants into each country every five years. Third, I examine the effects of policies incentivising and disincentivising fertility using the high and low variants of the UN population forecasts. Fourth, I simulate the effect of increasing the retirement age to 69, 74, and 79 years. The results of each of these experiments are discussed below.

6.1 Natural changes in age diversity

As explained above, using the medium variant of the UN's population forecasts for 2020-2100, we can calculate the GDP per capita predicted by age diversity leaving everything else constant for 2020-2100. This tells us whether changing age diversity will put on positive or negative pressures on GDP per capita.

Figure C.10 shows the effect of changing age diversity on income per capita by World Bank income status. Relative to 2015, by and large all regions will experience a downward pressure on income per capita. High- and upper-middle-income countries will experience some positive pressure in the middle of the century, and low-income countries will not experience any negative pressure until the middle of the century.

Figure C.11 shows the same graphs but broken down by World Bank region. There is somewhat more heterogeneity across regions now. North America will experience considerable positive pressure on GDP per capita until 2035, and then some negative pressure followed by a levelling off at a higher income per capita than in 2015. After a brief period of positive pressure until around 2030, the East Asia and Pacific, Europe and Central Asia, and Middle East and North Africa regions are predicted to experience cycles of negative and positive pressure. The positive pressure is expected to be stronger in the richer region of Europe and Central Asia. Latin American and the Caribbean, and South Asia will experience essentially uninterrupted negative pressure throughout the century. Finally, Sub-Saharan Africa is predicted to experience slight positive pressure until around 2040, followed by a decline.

The quantitative effects are not huge, but nevertheless significant. North America's GDP per capita, for instance, is predicted to be 4.1% higher by 2035 relative to 2015 due to pressures from changing age diversity. Meanwhile, South Asia's income per capita is predicted to be 5.3% lower in 2100 relative to 2015 due to negative pressure from changing age diversity.

These projections show that if the medium variant of the UN's population projections are realised, then age diversity will exert considerable pressure on GDP per capita around the globe. In many cases, the pressure is negative suggesting that policy interventions affecting age diversity in a desirable way can be useful for economic growth. I now discuss three policies that would affect age diversity.

6.2 Immigration policy

I simulate a simple immigration policy to evaluate immigration's effects in a set of countries. I assume that a given country lets in only 20-24-year-old immigrants, and they do so every five years. In every five-year period they let in a number of immigrants corresponding to 10% of the country's domestically-born 20-24-year-old population.

For example, if in 2020 the number of 20-24-year-olds predicted by the UN medium variant population projections is 100 people, then this number is modified to 110 people leaving the size of all other age groups unchanged. In 2025, the number of 20-24-year-olds is modified in the same fashion, but then the number of 25-29-year-olds is also incremented by 10% – effectively assuming that the immigrants from 2020 stayed in the country and aged accordingly.

Using these modified population projections, I once again calculate age diversity and its effect on GDP per capita. Figure C.12 shows the results of this exercise for four countries. In ageing, high-income countries such as Germany, Japan, and the United States, immigration altogether improves the predicted effects of age diversity to an extent. In 2100, German income per capita is predicted to be 0.8% higher, Japan's 0.7% higher, and that of the US 1.1% higher compared to the no-immigration case. Most of these gains would accrue later in the century, as the benefits of immigration would be small and even slightly negative until up to 2035.

On the other hand, the developing nation of Ethiopia would have a decidedly negative pressure on its income per capita from increased immigration until 2040. From 2040 onward, immigration would ultimately have zero effect on age diversity and thus income per capita in Ethiopia.

This analysis shows that controlling immigrant flows is an important policy parameter for managing age diversity, and that the age diversity impacts of immigration policy should be taken into account. While the economic effects of immigration through the channel of age diversity may be second-order, it is worth analysing what age groups a country's immigration policy should target in what decades. This would help determine the optimal composition of immigrants in a given country in a given period.

6.3 Fertility incentives

The United Nations presents its population projections for three fertility variants: low, medium, and high. In all the preceding analysis, the medium variant was used. However, an important policy tool that would affect age diversity in the long run is fertility incentives. To simulate the effect of a policy that incentivises or disincentivises fertility relative to its "natural" level, I calculate age diversity using projections from the UN's high and low variants.

Figure C.13 shows the effects for four countries. It appears that in the three developed countries of Germany, Japan, and the United States, lower fertility rates would be more advantageous for income per capita if we only look at the channel of age diversity. However, briefly in the 2070s and towards the end of the century the high-fertility scenarios start to look more favourable. Of course, any changes in fertility would take at least 20 years to affect the age diversity of the working-age population, so all three fertility variants produce the same pressures on income per capita until 2035. In the developing country of Ethiopia, fertility policies seem to matter little: for the 2060-2080 period fertility incentives are predicted to produce some extra growth via age diversity, while after 2080, this reverses and low fertility is more desirable.

These results underline that fertility incentives are an important policy tool for affecting income capita through the channel of age diversity. Fertility incentives and disincentive can both be useful depending on the country and time period in question.

6.4 Retirement age

In all previous policy experiments, I resorted my attention to the 20-64-year-old population effectively assuming that (most) people retire at age 64. I now consider what happens if the retirement age is pushed out to 69, 74, and even 79 years. I go back to the medium variant of fertility from the UN, but now expand the working-age population to include age groups older than 64.

Figure C.14 shows that starting around 2040, all four countries considered (Germany, Ethiopia, Japan, and the United States) are better off with a lower retirement age. This result is not entirely surprising as a higher retirement age increases age diversity considerably, while age diversity in all of these countries ultimately goes above the optimal level by the middle of the century. Increasing the retirement age somewhat, however, can be beneficial for the period before 2040.

Of course, just as with the other policy suggestions, I do not claim that these are the only impacts of retirement age policy. Nor do I claim that these are the strongest or even first-order effects. It is very well possible that due to other reasons, a higher retirement age would increase income per capita overall. I am merely showing that via the channel of age diversity, the pressure on GDP per capita would be negative beyond 2040 if retirement age is raised.

7 Discussion

In this paper, I document a new dimension via which demographics affects economics. I introduce the concept of age diversity, which is shown to be distinct from other measures of demographics. The key finding is that age diversity has a hump-shaped relationship with income per capita at the country- and European region-level. This non-linear relationship represents a trade-off. On the one hand, more age diversity is beneficial because it allows economies to take advantage of the different but complementary skills of different age groups. On the other hand, age diversity is costly, because not all age groups' skills are in equal demand in an economy in general. The optimal level of age diversity is thus pinned down by the relative importance of age-specific skills.

This study underlines the importance of demographic factors in the growth process, and suggests several factors policy-makers should pay attention to. First, it is shown that age diversity will be changing naturally over the 21st century, and economically damaging shifts can potentially be addressed through policy. Furthermore, the paper argues that immigration, fertility incentives, and the retirement age are key policy tools which have a substantial impact on age diversity. The age diversity channel of these policies can reinforce or offset the first-order effects of these policies.

The findings also suggest that not all is doomed if a government's hands are tied in managing age diversity. In particular, considerable variaton in the optimal level of age diversity is documented between different sectors. Therefore, ensuring that the sectoral composition of an economy is consistent with its demographic structure can move an economy's optimal age diversity closer to its actual one. This implies that regional and national policymakers should both pay attention to demographic projections, and incentivise the emergence of industries that can thrive in their particular demographic environment in the long run. In other words, similarly to Gu and Stoyanov (2019), this study emphasises the connection between the comparative advantage of an economy and its age structure.

Finally, based on the findings of this paper, it is conceivable that schemes that keep older workers' skills up-to-date in a rapidly changing environment can raise the optimal level of age diversity. Things such as continuous education schemes can ensure that the education gap between young and old workers is not as large as it would be in the absence of such interventions. This would make old workers better substitutes for young workers than they otherwise would be.

A variety of questions about age diversity remain open to future research. Perhaps the most important question would be further research into the determinants of the optimal level of age diversity. This would be a crucial parameter for understanding how policy changes that will only show their effects decades in the future will impact economic output through age diversity. Another question is how age diversity and its economic impacts interact with changes in other demographic indicators. This is especially important, because if age diversity is used as a policy tool, policy-makers will inadvertently impact other demographic characteristics as well when they try to manage age diversity.

References

Aiyar, S., Ebeke, C., and Xiaobo, S. (2016). The impact of workforce aging on european productivity. *IMF Working Papers 16/238*.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–94.

Alesina, A., Harnoss, J., and Rapaport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–38.

Ang, J. and Madsen, J. (2015). Imitation versus innovation in an aging society: International evidence since 1870. *Journal of Population Economics*, 28(2):299–327.

Arntz, M. and Gregory, T. (2014). What old stagers could teach us - examining age complementarities in regional innovation systems. *Working paper*.

Ashraf, Q. and Galor, O. (2013). The Out of Africa hypothesis, human genetic diversity and comparative development. *American Economic Review*, 103(1):1–46.

Auerbach, A. and Kotlikoff, L. (1987). *Dynamic fiscal policy*. Cambridge University Press, Cambridge, MA.

Backes-Gellner, U. and Veen, S. (2013). Positive effects of ageing and age diversity in innovative companies large-scale empirical evidence on company productivity. *Human Resource Management Journal*, 23(3):279–95.

Burke, D. and Mackay, D. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society B*, 352(1363):1845–56.

Caselli, R., Dueck, A., Osborne, D., Sabbagh, M., Connor, D., Ahern, G., Baxter, L., Rapcsak, S., Shi, J., Woodruff, B., Locke, D., Hoffman Snyder, C., Alexander, G., Rademakers, R., and Reiman, E. (2009). Longitudinal modeling of age-related memory decline and the apoe e4 effect. *The New England Journal of Medicine*, 361:255–63.

Cohen, D. and Leker, L. (2014). Health and education: Another look with the proper data. *CEPR Discussion Papers 9940*.

Derrien, F., Kecskes, A., and Nguyen, P. (2018). Labor force demographics and corporate innovation. *Working paper*.

Feenstra, R., Inklaar, R., and Timmer, M. (2015). The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–82.

Feyrer, J. (2007). Demographics and productivity. *The Review of Economics and Statistics*, 89(1):100–09.

Freeman, R. (1979). The effect of demographic factors on age-earnings profiles. *Journal of Human Resources*, 14(3):289–318.

Gregory, T. and Patuelli, R. (2015). Demographic ageing and the polarization of regions an exploratory spacetime analysis. *Environment and Planning A*, 47(5):1192–210.

Gu, K. and Stoyanov, A. (2019). Skills, population aging, and the pattern of trade. *Review of International Economics*, pages 1–21.

Guest, R. and Shacklock, K. (2005). The impending shift to an older mix of workers: Perspectives from the management and economics literatures. *International Journal of Organisational Behaviour*, 10(3):713–28.

Hall, C., Lipton, R., Sliwinski, M., Katz, M., Derby, C., and Verghese, J. (2009). Cognitive activities delay onset of memory decline in persons who develop dementia. *Neurology*, 73(5):356–61.

Hartshorne, J. and Germine, L. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26(4):433–43.

Hjort, J. (2014). Ethnic divisions and production in firms. *Quaterly Journal of Economics*, 129(4):1899–946.

Irmen, A. and Litina, A. (2016). Population aging and inventive activity. *CESifo Working Paper Series No. 5841*.

Karlgaard, R. (2019). *Late Bloomers: The Power of Patience in a World Obsessed with Early Achievement*. Currency, New York, NY.

Keilman, N. (1998). How accurate are the United Nations World Population Projections? *Population and Development Review*, 24:15–41.

Kremer, M. and Thomson, J. (1998). Why isn't convergence instantaneous? young workers, old workers, and gradual adjustment. *Journal of Economic Growth*, 3(1):5–28.

Kuchler, H. (2017). Silicon Valley ageism: They were, like, wow, you use Twitter?. *Financial Times*. 30 July.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and T., S. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849.

Laitner, J. (1990). Tax changes and phase diagrams for an overlapping generations model. *Journal of Political Economy*, 98(1):193–220.

Lam, D. (1989). Population growth, age structure, and age-specific productivity: does a uniform age distribution minimize lifetime wages? *Journal of Population Economics*, 2(3):189–210.

Lee, J. and Lee, H. (2016). Human capital in the long run. *Journal of Development Economics*, 122:147–69.

Mukherjee, R. (2017). Report: Ageism in the tech industry. *indeed blog*. 19 October.

Murphy, K., Plant, M., and Welch, F. (1988). Cohort size and earnings in the united states. In Lee, R., Aurthur, W., and Rodgers, G., editors, *Economics of changing age distributions in developed countries*. Clarendon Press, Oxford.

Pelisson, A. and Hartmans, A. (2017). The average age of employees at all the top tech companies, in one chart. *Business Insider*. 11 September.

Penn World Table (2014). Human capital in PWT 9.0. https://www.rug.nl/ggdc/docs/human_capital_in_pwt_90.pdf. Accessed: 2019-02-28.

Prskawetz, A., Fent, T., and Guest, R. (2008). Workforce aging and labor productivity: The role of supply and demand for labor in the G7 countries. *Population and Development Review*, 34:298–323.

Prskawetz, A., Tsachev, T., and Veliov, V. (2012). Optimal education in an age-structured model under changing labor demand and supply. *Macroeconomic Dynamics*, 16(2):159–83.

Psacharopoulos, G. and Patrinos, A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, 26(5):445–58.

Rojas, J. (2005). Life-cycle earnings, cohort size effects and social security: a quantitative exploration. *Journal of Public Economics*, 89(2-3):465–85.

Roundtree, L. (2011). Age: A 21st century diversity imperative. https://www.bc.edu/research/agingandwork/archive_pubs/case_AgeDiversity.html. Accessed: 2018-03-16.

Salthouse, T. (2003). Memory aging from 18 to 80. *Alzheimer Disease & Associated Disorders*, 17(3):162–7.

Samuelson, P. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–46.

United Nations (1985). *World Population Prospects: Estimates and Projections As Assessed in 1982*. New York, NY.

United Nations (2017). Population by age groups - both sexes. https://esa.un.org/unpd/wpp/Download/Standard/Population/. Accessed: 2018-03-19.

Weil, D. (1997). The economics of population aging. In Rosenzweig, M. and Stark, O., editors, *Handbook of Population and Family Economics*, pages 967–1014. North Holland.

Welch, F. (1979). Effect of cohort size on earnings: the baby boom babies' financial bust. *Journal of Political Economy*, 87(5):565–98.

Whitley, E., Deary, I., Ritchie, S., Batty, D., Kumari, M., and Benzeval, M. (2016). Variations in cognitive abilities across the life course: Cross-sectional evidence from understanding society: The uk household longitudinal study. *Intelligence*, 59:39–50.

A Proofs

A.1 Proof of Lemma 1

The size of each group is the solution to the following system of three equations in three unknowns,

$$
\begin{aligned}
n_0+n_1+n_2 &= 1\\
27n_0+42n_1+57n_2 &= \mu\\
n_0(27-\mu)^2+n_1(42-\mu)^2+n_2(57-\mu)^2 &= \sigma^2,
\end{aligned}
$$

where $n_0, n_1, n_2 \in [0,1]$ are the shares of each group, and μ and σ are given. An underlying assumption behind this result is that within each age group, people are uniformly distributed in terms of age. So the mean age of the 20-34-year-old group is 27, and so on for the other groups.

A.2 Proof of Proposition 1

Using the production function and the aggregate factor quantities from (1)-(2), we have that

$$
\begin{aligned}
\ln Y = \alpha \ln\left[h_0\left(1-\delta\left(\frac{\sigma^2}{225}+\frac{1458}{450}+\frac{\mu^2}{225}+\frac{108}{450}\mu\right)\right)\right]+\\
(1-\alpha)\ln\left[e_0\left(1+\varepsilon\left(\frac{\sigma^2}{225}+\frac{1458}{450}+\frac{\mu^2}{225}+\frac{108}{450}\mu\right)\right)\right].
\end{aligned}
$$

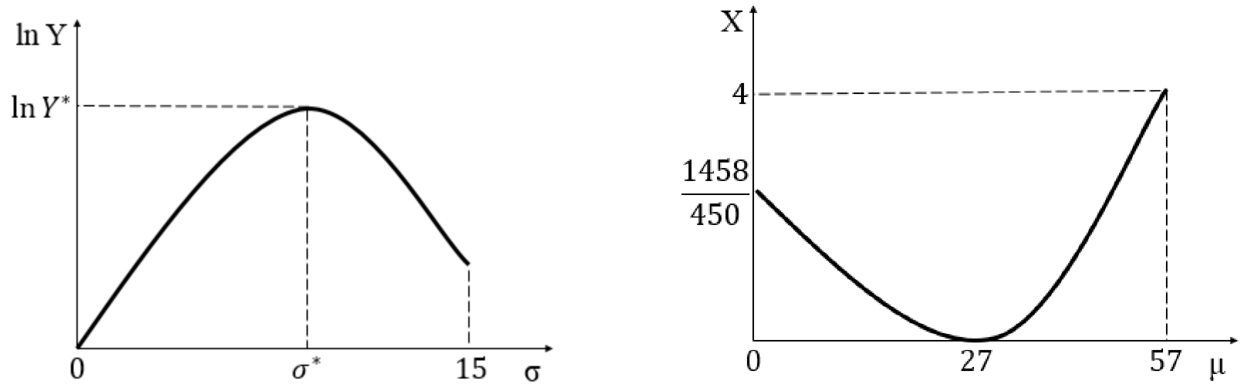
The derivative of this expression with respect to σ is

$$
\frac{\partial \ln Y}{\partial \sigma}=\frac{2}{225}\sigma\left[(1-\alpha)\frac{\varepsilon}{1+\varepsilon\left(\frac{\sigma^2}{225}+X\right)}-\alpha\frac{\delta}{1-\delta\left(\frac{\sigma^2}{225}+X\right)}\right],
$$

where $X \equiv \frac{1458}{450}+\frac{\mu^2}{225}-\frac{108}{450}\mu$. To show this derivative implies a hump-shaped relationship between $\ln Y$ and σ, note that the derivative is positive if and only if σ is below a threshold σ^*. And if $\sigma > \sigma^*$ then the derivative is negative. Therefore, it is a hump-shaped curve. This is illustrated graphically in Figure A.7a. More formally, $\frac{\partial \ln Y}{\partial \sigma} > 0$ if and only if

$$
(\sigma^*)^2 \equiv 225\left[\frac{1-\alpha}{\delta}-\frac{\alpha}{\varepsilon}-X\right] > \sigma^2, \tag{7}
$$

and $\frac{\partial \ln Y}{\partial \sigma} < 0$ if $\sigma > \sigma^*$. In order for the threshold to be binding, we need it to be positive and not too high. The minimum attainable age diversity in this model is $\sigma = 0$, which happens when all people are in the same age group. So as long as $\sigma^* > 0$, there is at least some region of σ where increasing diversity leads to higher output. The maximum attainable age diversity in this model is $\sigma = 15$, which happens when $n_0 = n_2 = 0.5$ and $n_1 = 0$. So as long as $\sigma^* < 15$, there is at least some region of σ where increasing diversity leads to



(a) Output and diversity

(b) Behaviour of X

Figure A.7: Visual representation of the proof of Proposition 1

lower output.

It is easily shown from (7) that $\sigma^* > 0$ if and only if $\alpha < \frac{\varepsilon - X\delta\varepsilon}{\varepsilon+\delta}$. Similarly, we can also show that $\sigma^* < 15$ if and only if $\alpha > \frac{\varepsilon-\delta\varepsilon-X\delta\varepsilon}{\varepsilon+\delta}$. Both of these bounds on α hold by Assumption 1. It is also apparent that, as required, the upper bound on α is higher than the lower bound.

The last important thing to show that the required upper bound on α is positive, and that the required lower bound on α is less than 1. Otherwise, we cannot get a hump-shape for $\alpha \in (0,1)$. The upper bound is positive if and only if $\varepsilon - X\delta\varepsilon > 0$. This happens when $X\delta < 1$. The lower bound is less than 1 if and only if $\varepsilon - \delta\varepsilon - X\delta\varepsilon < \varepsilon + \delta$. This happens when $-\varepsilon(1+X) < 1$. Both of these conditions hold if $\delta \in \left[0, \frac{1}{4}\right)$, which is true by Assumption 1.

To see why, we need to look at what values $X = \frac{1458}{450} + \frac{\mu^2}{225} - \frac{108}{450}\mu$ can take. First, note that X is a U-shaped function of μ with a unique minimum, which is attained at $\mu = 27$ at which point $X = 0$. Second, note that in our model it must be that $\mu \in [27, 57]$. Combining these two facts, it is clear after plugging in for μ using its lower and upper bounds that $X \in [0,4]$. It is apparent then that $X\delta < 1$ because $\delta < \frac{1}{4}$, and that $-\varepsilon(1+X) < 1$, because $\varepsilon, X > 0$. This reasoning is illustrated graphically in Figure A.7b.

A.3 Proof of Proposition 2

Following the proof of Proposition 1, it is clear that the diversity that maximises output is given by σ^* in Equation (7). It is apparent by observation that this is decreasing in α.

B Description of model solution

This appendix describes how the model is solved. The algorithm we use to solve for the transition between the initial and final steady state as shown in Table 1 uses the following procedure (Auerbach and Kotlikoff (1987)).

1. Solve the model for the initial ($t = 0$) and final ($t = F$) steady states that may differ in parameter values, though we shall assume they are identical.

2. Guess a sequence of human capital wage rates for the transition path, $\{w_t^h\}_{t=0}^F$ by linearly interpolating between w_0^h and w_F^h.

3. In each iteration i for a sequence of wage rates $\{w_{t,i}^h\}_{t=1}^{F-1}$, do the following.

   (a) Calculate $w_{t,i}^e = (1-\alpha_t)\left(\frac{\alpha_t}{w_{t,i}^h}\right)^{\frac{\alpha_t}{1-\alpha_t}}$ for each $t \in [1, F-1]$. This expression follows from combining the firms' first-order conditions.

   (b) Given $\{w_{t,i}^h\}_{t=1}^{F-1}$ and $\{w_{t,i}^e\}_{t=1}^{F-1}$, solve the household optimisation problem in (3) in all periods $t \in [1, F-1]$.

   (c) Aggregate across all households to calculate the aggregate human capital stock, $H_t = \sum_{k=0}^T N_{k,t} h_{k,t}$ for $t \in [1, F-1]$. Calculate aggregate experience as $E_t = \sum_{k=0}^T N_{k,t} e_{k,t}$.

   (d) Calculate $\tilde{w}_{t,i}^h = \alpha_t \left(\frac{H_t}{E_t}\right)^{\alpha_t - 1}$ for $t \in [1, F-1]$. This is the human capital wage rate from the firms' first-order conditions.

   (e) Calculate $\sum_{t=1}^{F-1} |w_{t,i}^h - \tilde{w}_{t,i}^h|$ as the discrepancy between the initial wage sequence guess and the one implied by the FOCs. If the gap is smaller than a tolerance parameter (0.02), then stop. Otherwise, make a new guess for the wage sequence, $w_{t,i+1}^h = \omega w_{t,i}^h + (1-\omega)\tilde{w}_{t,i}^h$ for $t \in [1, F-1]$ and a dampening parameter $\omega \in (0,1)$, and go back to Step 3. We pick $\omega = 0.5$.

Two further algorithms need to be described in order to fully characterise the solution. First, it needs to be clarified how we solve for the initial and final steady states in Step 1. Second, it needs to be clarified how we solve the household model in Step 3b. We now explain these two procedures.

The solution procedure for the steady-state values w_0^h and w_F^h is essentially identical to the out-of-steady-state solution procedure (Auerbach and Kotlikoff (1987)). We go through the following steps.

1. Guess a steady-state human capital wage rate w^h.

2. In each iteration i for a wage rate w_i^h, do the following.

   (a) Calculate $w_i^e = (1-\alpha)\left(\frac{\alpha}{w_i^h}\right)^{\frac{\alpha}{1-\alpha}}$. This expression follows from combining the firms' first-order conditions.

   (b) Given w_i^h and w_i^e, solve the household optimisation problem in (3).

(c) Aggregate across all households to calculate the aggregate human capital stock, $H = \sum_{k=0}^{T} N_k h_k$. Calculate aggregate experience as $E = \sum_{k=0}^{T} N_k e_k$.

(d) Calculate $\tilde{w}_i^h = \alpha \left(\frac{H}{E}\right)^{\alpha-1}$. This is the human capital wage rate from the firms' first-order conditions.

(e) Calculate $|w_i^h - \tilde{w}_i^h|$ as the discrepancy between the initial wage guess and the one implied by the FOCs. If the gap is smaller than a tolerance parameter (0.01), then stop. Otherwise, make a new guess for the wage, $w_{i+1}^h = \omega w_i^h + (1-\omega)\tilde{w}_i^h$ for a dampening parameter $\omega \in (0,1)$, and go back to Step 2. We pick $\omega = 0.5$.

Finally, the household problem in Step 3b of the out-of-steady-state solution and Step 2b of the steady-state solution needs to be obtained. For this, we follow Samuelson (1969) and express the household problem with total wealth as the state variable, which allows us to solve for household consumption in every period as a fraction of total wealth.[12] This procedure works as follows. First, we change the timing convention in our budget constraint and rewrite it as

$$c_t + \frac{a_{t+1}}{1+r} = y_t + a_t$$
$$a_{t+1} = (a_t + y_t - c_t)(1+r), \tag{8}$$

where y_t denotes income in period t. In our model,

$$y_0 = w_0^h h_0 + w_0^e e_0 - \sigma h_0^2$$
$$y_t = w_t^h h_t + w_t^e e_t \text{ for } t \in [1, T].$$

Then, we define the present value of future income as $d_t = \sum_{s=t+1}^{T} y_t (1+r)^{(-s-t)}$. This can be written recursively as $d_{t+1} = d_t(1+r) - y_{t+1}$. If we add this to both sides of the rewritten budget constraint in (8), we get

$$a_{t+1} + d_{t+1} = (a_t + y_t - c_t)(1+r) + d_t(1+r) - y_{t+1}.$$

If we define the household's total wealth as $w_t = a_t + y_t + d_t$, then we can write the budget constraint as

$$w_{t+1} = (w_t - c_t)(1+r). \tag{9}$$

It is now possible to solve the dynamic programming problem

$$V_t(w_t) = \max_{c_t} \frac{c_t^{1-\theta}}{1-\theta} + \beta V_{t+1}(w_{t+1})$$

subject to (9). We begin by noting that $c_T = w_T$, which allows us to derive with backward induction that

[12] I would like to thank Alexander Ludwig whose lecture notes pointed this method out to me.

$c_t = m_t w_t$ in any period $t \in [0,T]$ with $m_t = \frac{1-b}{1-b^{T-t+1}}$ and $b = (\beta(1+r)^{1-\theta})^{\frac{1}{\theta}}$. Then the solution of the household problem follows the following procedure.

1. Create a grid of values for human capital choice $\{h_{0,i}\}_{i=0}^{m}$, where m denotes the number of grid points we consider. We pick $m = 500$ and our grid has increments of 0.01.

2. For each $h_{0,i}$ and given the wage rates w^h and w^e, solve the household problem by calculating $\{c_t\}_{t=0}^{T}$ using the formula $c_t = m_t w_t$. In each period, we then know that $a_{t+1} = (y_t + a_t - c_t)(1+r)$.

3. Once we iterated through all $h_{0,i}$, pick the $h_{0,i}$ that leads to the highest utility.

The solution method's validity is verified by checking that the budget constraint indeed holds in every period. Further, we also graphically verify that the human capital investment we pick is a seemingly unique maximum. This is illustrated in Figure C.2, which plots the utility that each h_0 we considered leads to for the initial steady state.

C Figures

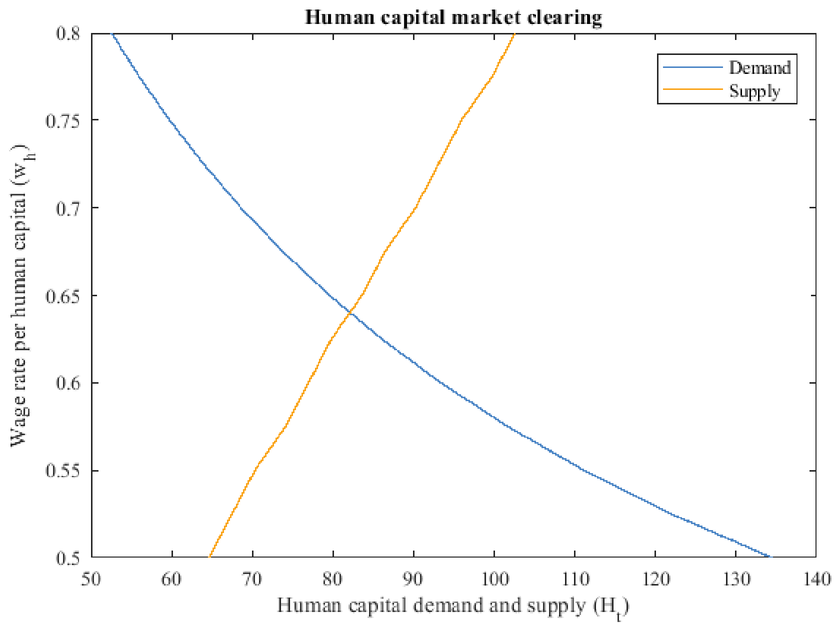


Figure C.1: Existence and uniqueness of market equilibrium in the model

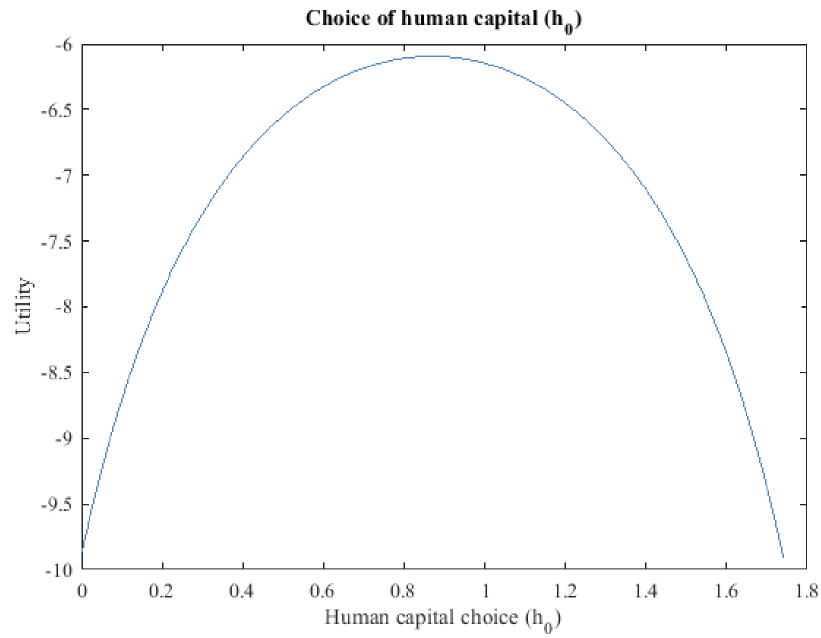


Figure C.2: Existence and uniqueness of utility-maximising h_0 in the model

Figure C.3: Life-cycle earnings profile in steady state in the model

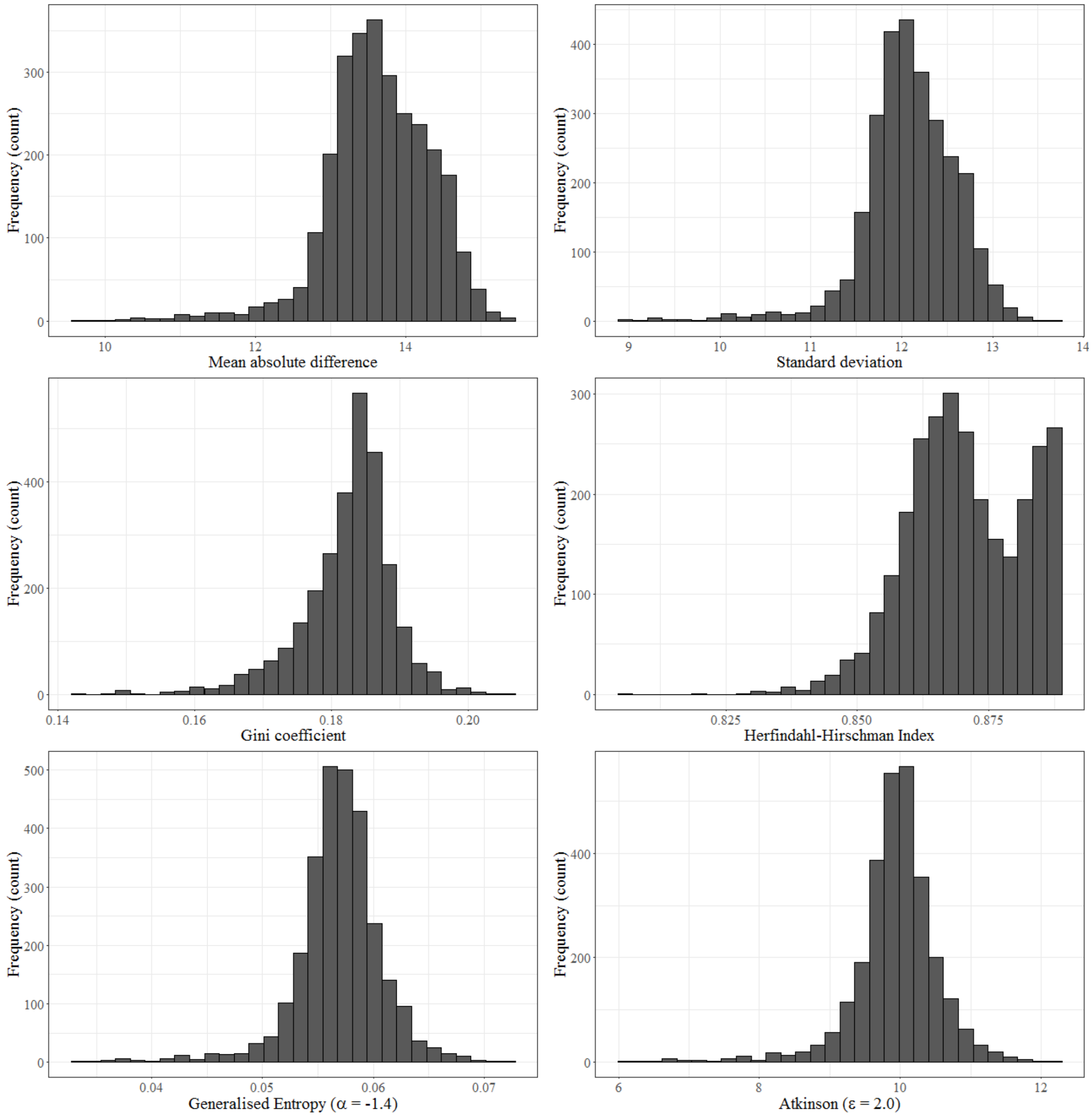


Figure C.4: Histograms of age diversity metrics

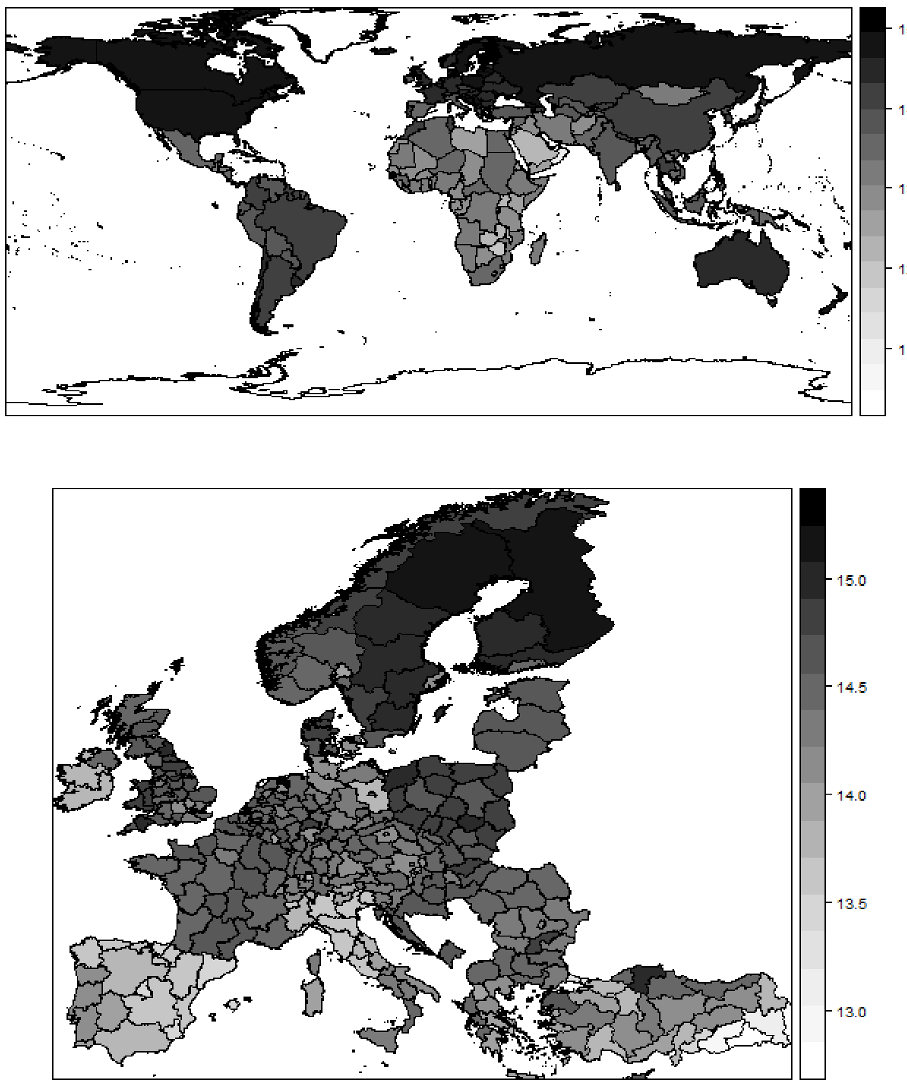


Figure C.5: Age diversity (MAD) at the country-level and in European regions (2015)

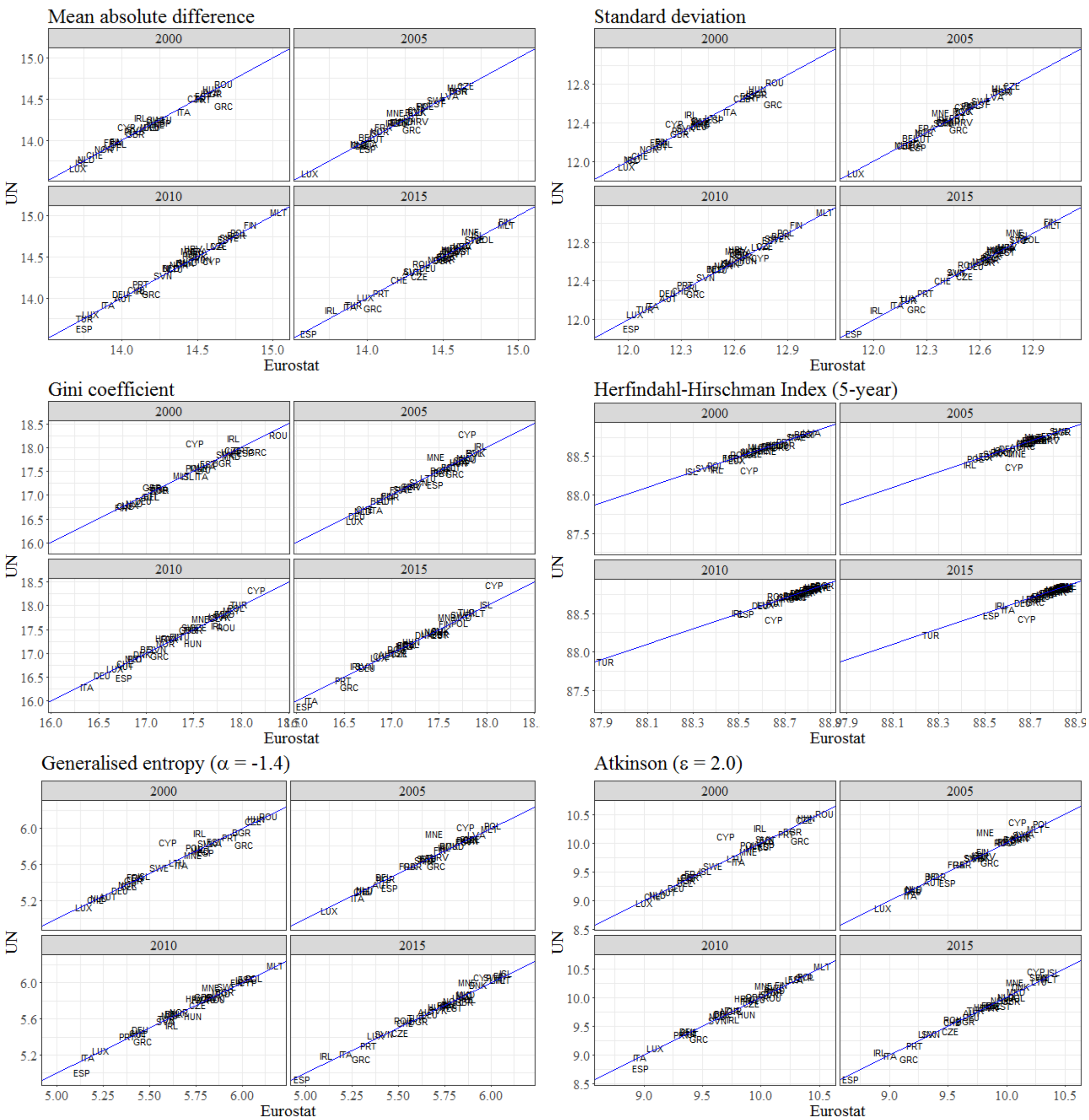


Figure C.6: Cross-check between the UN and Eurostat data

Note: These plots shows that our calculation of age diversity using UN data is consistent with our calculation of age diversity using Eurostat data. The figures plot the country-level measure of age diversity calculated from UN data on the vertical axis against the measure of age diversity calculated from the Eurostat data (regional data aggregated to the country level) on the horizontal axis. The lines in the figures are 45-degree lines. The plots are shown for six different measures of age diversity.

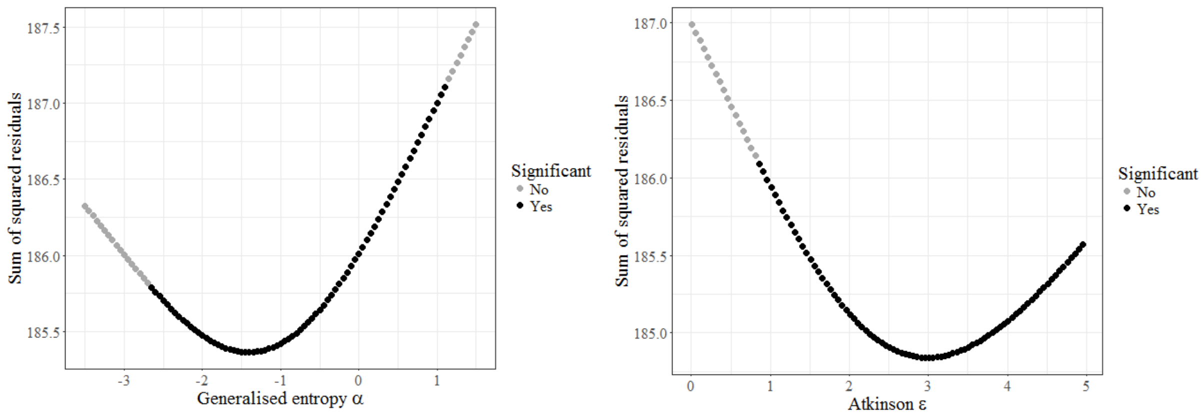


Figure C.7: Finding the optimal parameters for the GE and Atkinson indices (country-level)

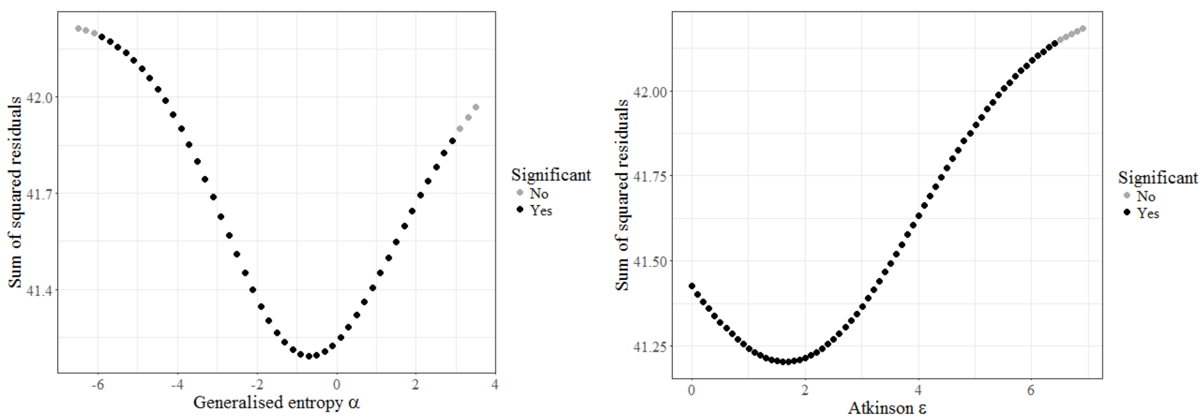


Figure C.8: Finding the optimal parameters for the GE and Atkinson indices (European regions)

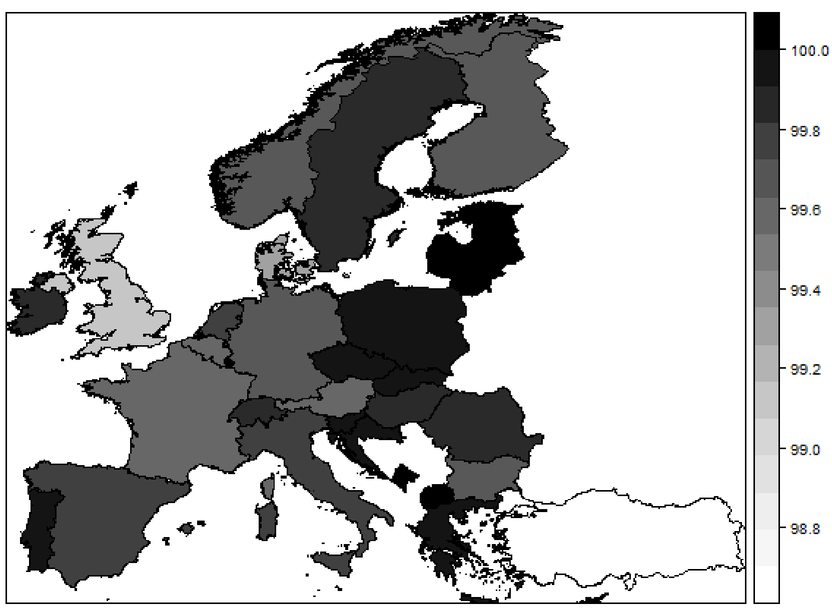


Figure C.9: Percentage of total age diversity explained by within-region age diversity (2015)

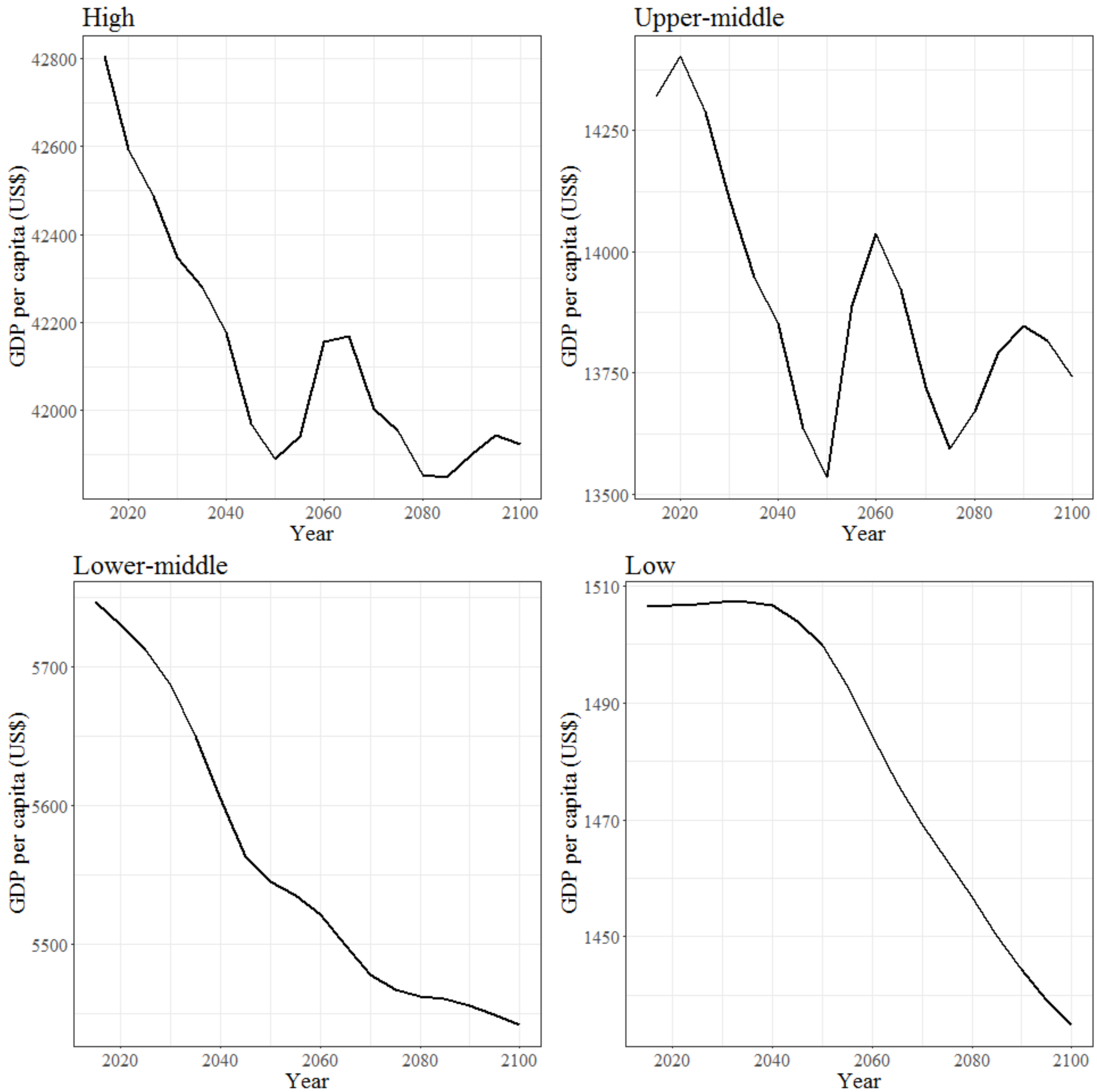


Figure C.10: Predicted effect of age diversity on GDP per capita by World Bank income status

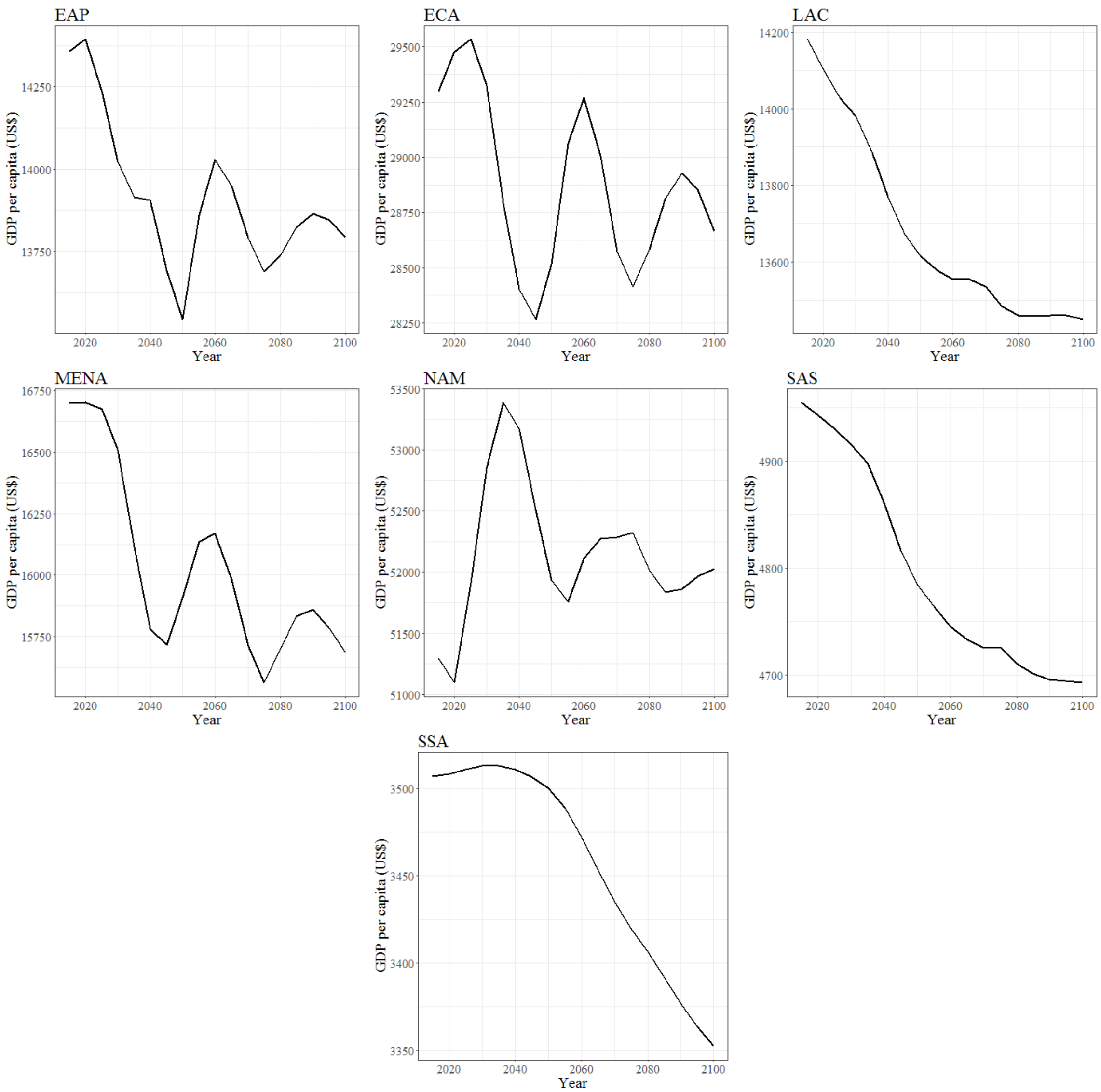


Figure C.11: Predicted effect of age diversity on GDP per capita by World Bank region

Note: “EAP” is East Asia and Pacific, “ECA” is Europe and Central Asia, “LAC” is Latin American and the Caribbean, “MENA” is Middle East and North Africa, “NAM” is North America, “SAS” is South Asia, and “SSA” is Sub-Saharan Africa.

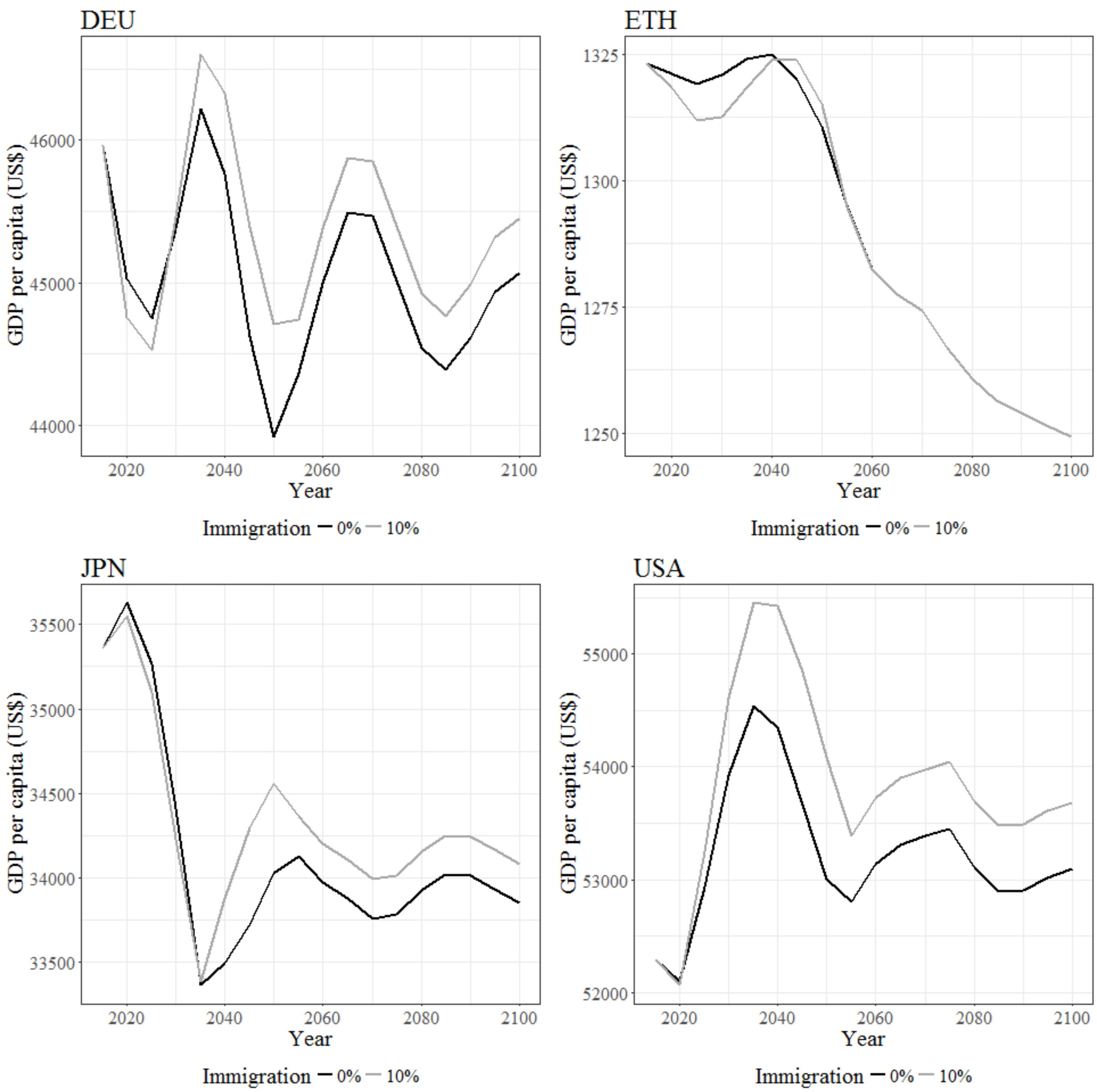


Figure C.12: Predicted effect of increased immigration for four countries

Note: "DEU" is Germany, "ETH" is Ethiopia, "JPN" is Japan, and "USA" is the United States. This figure shows for four countries how increased immigration would influence the predicted effect of future changes in age diversity on GDP per capita.

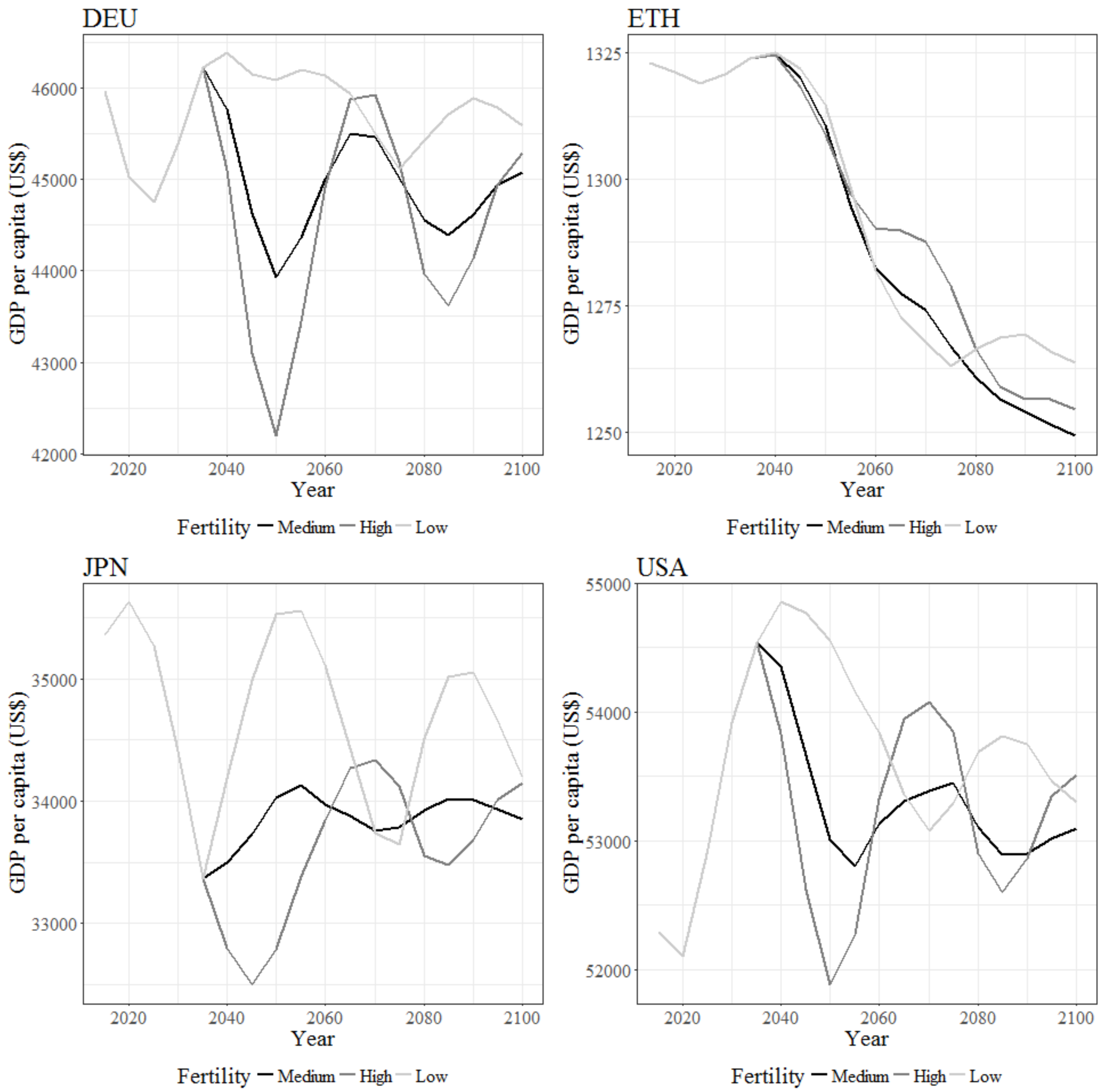


Figure C.13: Predicted effect of fertility incentives for four countries

Note: “DEU” is Germany, “ETH” is Ethiopia, “JPN” is Japan, and “USA” is the United States. This figure shows for four countries how different levels of fertility would influence the predicted effect of future changes in age diversity on GDP per capita.

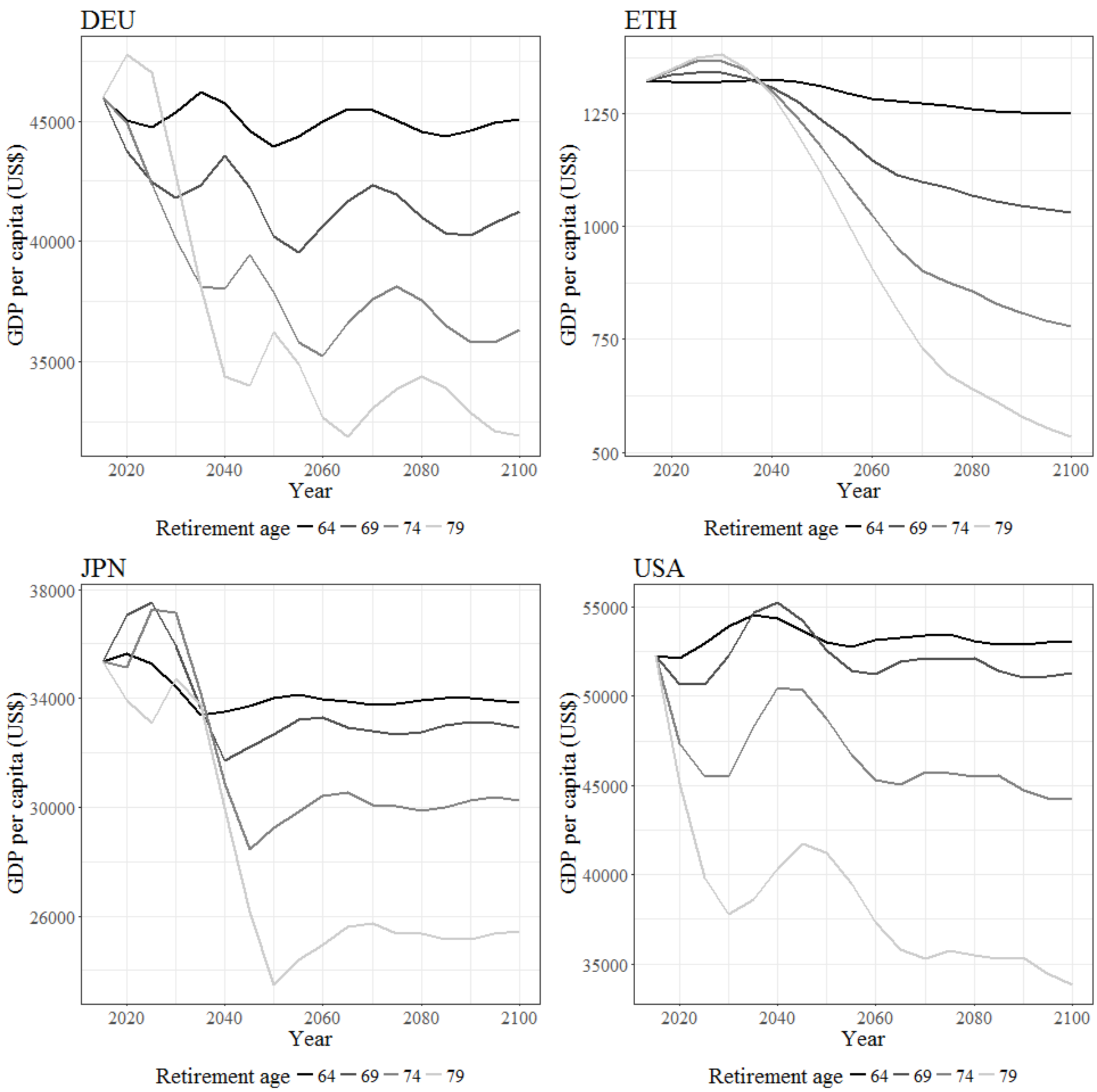


Figure C.14: Predicted effect of changes in the retirement age for four countries

Note: "DEU" is Germany, "ETH" is Ethiopia, "JPN" is Japan, and "USA" is the United States. This figure shows for four countries how different retirement ages would influence the predicted effect of future changes in age diversity on GDP per capita.

D Tables

Table D.1: Parameter values used for model simulation

Param.	Name	Value	Explanation
θ	Risk-aversion	2	Intermediate level of risk aversion.
σ	Cost of human capital	1	Normalised to 1.
β	Discount factor	0.8219	Corresponds to a discount rate of 4% per annum.
ε	Growth of experience	0.0357	Calibrated to match concave life-cycle earnings profile (Lagakos et al., 2018).
k	Convexity of experience	1.5	Calibrated to match concave life-cycle earnings profile (Lagakos et al., 2018).
δ	Depreciation of human capital	0.0016	Corresponds to a total loss of 10% of human capital over the life-cycle.
e_0	Initial level of experience	1	Normalised to 1.
N	Steady-state population	100	Normalised to 100.
α	Returns to education	0.82	Chosen to approximately match the empirically-observed optimal age diversity.
r	Saving/borrowing interest rate	0.65	Clears the asset market in the initial steady state. Corresponds to 10.5% per annum.

Table D.2: Summary of our measures of age diversity

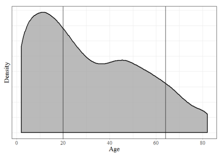
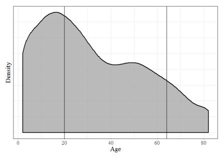
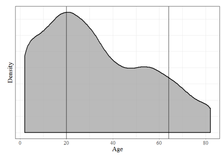
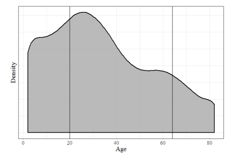
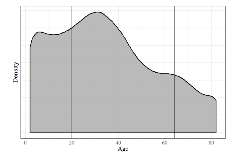
Name	Formula	Type	Notes
Mean absolute difference (MAD)	$\sum_i \sum_j p_i p_j \lvert a_i - a_j \rvert$	Non-categorical	Expected age difference between two randomly picked people.
Standard deviation (SD)	$\sum_i p_i (a_i - \bar{a})^2$	Non-categorical	-
Gini coefficient	$\frac{1}{2\sum_i p_i a_i} \sum_i \sum_j p_i p_j \lvert a_i - a_j \rvert$	Non-categorical	-
Herfindahl-Hirschman Index (HHI)	$1 - \sum_i p_i^2$	Categorical	Probability that two randomly picked people are from a different age group.
Generalised Entropy Index (GE)	$\frac{1}{\alpha(\alpha-1)} \sum_i p_i [(a_i/\bar{a})^\alpha - 1]$ for $\alpha \neq 0, 1$ $\sum_i p_i \frac{a_i}{\bar{a}} \ln \frac{a_i}{\bar{a}}$ for $\alpha = 1$ $\sum_i p_i \ln \frac{a_i}{\bar{a}}$ for $\alpha = 0$	Non-categorical	Includes mean log deviation, Theil index, coefficient of variation as special cases.
Atkinson Index	$1 - \frac{1}{\bar{a}} (\sum_i p_i a_i^{1-\varepsilon})^{1/(1-\varepsilon)}$ for $0 \leq \varepsilon \neq 1$ $1 - \frac{1}{\bar{a}} \prod_i p_i a_i$ for $\varepsilon = 1$	Non-categorical	-

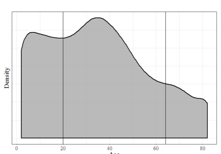
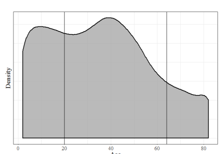
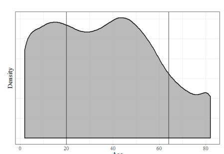
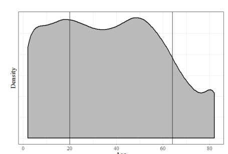
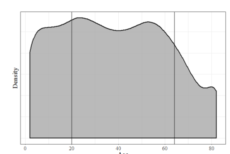
Note: Throughout this table i is the index of an age group, a_i is the (mean) age of group i, p_i is the share of group i in the working-age population, and $\bar{a}$ is the mean age of the working-age population. All other symbols are parameters of specific measures.

Table D.3: Correlation matrix of age diversity metrics

	MAD	SD	Gini	HHI	GE	Atkinson
MAD	1	0.985	0.330	0.793	0.852	0.784
SD	0.985	1	0.466	0.683	0.913	0.862
Gini	0.330	0.466	1	-0.242	0.750	0.829
HHI	0.793	0.683	-0.242	1	0.388	0.278
GE	0.852	0.913	0.750	0.388	1	0.991
Atkinson	0.784	0.862	0.829	0.278	0.991	1

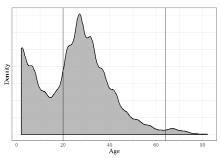
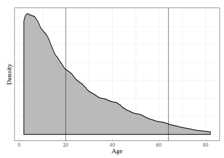
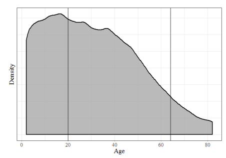
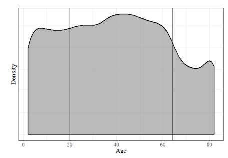
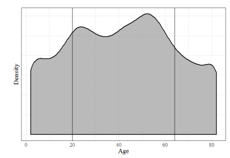
Table D.4: Population structure and age diversity in the United States, 1970-2015

	1970	1975	1980	1985	1990
					
MAD	14.797	14.923	14.805	14.377	13.834
SD	12.918	13.069	13.021	12.678	12.182
Gini	0.185	0.190	0.191	0.187	0.178
HHI	0.886	0.883	0.880	0.880	0.881
GE	0.063	0.065	0.064	0.060	0.055
Atkinson	10.789	11.065	10.957	10.386	9.622

	1995	2000	2005	2010	2015
					
MAD	13.513	13.663	14.087	14.563	14.873
SD	11.855	11.942	12.288	12.694	12.963
Gini	0.172	0.170	0.172	0.175	0.179
HHI	0.882	0.884	0.887	0.888	0.889
GE	0.052	0.053	0.056	0.059	0.061
Atkinson	9.142	9.239	9.673	10.140	10.470

Note: This table shows how population structure and age diversity varied in the US from 1970 to 2015. The graphs show the population's age distribution. The two vertical lines in each graph are at age 20 and 64, and they denote the bounds of the working-age population.

Table D.5: Population structure and age diversity in various countries and years

					
	ARE 1980	SLB 1985	BHS 2005	FRA 2010	LTU 2015
MAD	9.694	13.246	13.630	14.578	14.709
SD	8.982	11.834	11.962	12.706	12.827
Gini	0.149	0.185	0.176	0.173	0.175
HHI	0.819	0.862	0.880	0.889	0.888
GE	0.035	0.056	0.055	0.058	0.060
Atkinson	6.547	9.914	9.554	9.951	10.281

Note: This table shows how different population structures are associated with different levels of age diversity. Each column is for a different country/year pair going left to right from lowest MAD to high MAD. The countries are “ARE” for United Arab Emirates, “SLB” for the Solomon Islands, “BHS” for the Bahamas, “FRA” for France, and “LTU” for Lithuania. The graphs show the population’s age distribution. The two vertical lines in each graph are at age 20 and 64, and they denote the bounds of the working-age population.

Table D.6: Age diversity and GDP per capita (levels)

	Dependent variable:				
	Log GDP per capita				
	(1)	(2)	(3)	(4)	(5)
Age diversity	2.989***	2.555***	2.225**	2.811***	2.614***
	(0.814)	(0.816)	(0.882)	(0.835)	(0.883)
Age diversity sq.	$-$0.108***	$-$0.095***	$-$0.084***	$-$0.112***	$-$0.103***
	(0.029)	(0.029)	(0.032)	(0.031)	(0.033)
Young-age dependency		$-$1.096***	$-$0.310		0.119
		(0.151)	(0.274)		(0.377)
Old-age dependency		4.219***	0.703		$-$1.340
		(0.846)	(1.484)		(1.652)
Mean age			0.090***		0.158***
			(0.034)		(0.047)
Fraction 20-29				$-$8.435***	$-$0.785
				(2.673)	(2.460)
Fraction 30-39				$-$7.160**	$-$2.723
				(3.557)	(3.240)
Fraction 40-49				$-$7.910**	$-$4.564
				(3.180)	(2.868)
Fraction 50-59				$-$2.177	$-$1.294
				(2.339)	(1.906)
Peak	13.809	13.459	13.178	12.512	12.681
	(0.068)	(0.094)	(0.173)	(0.365)	(0.324)
Observations	1,805	1,805	1,805	1,805	1,805
R^2	0.907	0.927	0.928	0.913	0.930
Adjusted R^2	0.897	0.918	0.920	0.903	0.922
Note:	*p<0.1; **p<0.05; ***p<0.01				

Note: This table shows a hump-shaped relationship between age diversity and GDP per capita. The specification regresses levels on levels. All columns include country and year fixed effects. Robust standard errors are clustered at the country-level.

Table D.7: Age diversity and GDP per capita (first-differenced)

	Dependent variable:				
	Log GDP per capita				
	(1)	(2)	(3)	(4)	(5)
Age diversity	0.973**	1.190**	1.150**	0.997**	1.229**
	(0.474)	(0.480)	(0.481)	(0.483)	(0.510)
Age diversity sq.	−0.034**	−0.041**	−0.041**	−0.038**	−0.046**
	(0.017)	(0.017)	(0.017)	(0.018)	(0.019)
Young-age dependency		−0.652***	−0.281*		−0.120
		(0.096)	(0.154)		(0.209)
Old-age dependency		0.618	−0.979		−1.248
		(0.680)	(0.928)		(1.022)
Mean age			0.055**		0.073***
			(0.023)		(0.028)
Fraction 20-29				−0.686	0.065
				(1.155)	(1.317)
Fraction 30-39				−1.277	−0.655
				(1.546)	(1.618)
Fraction 40-49				−1.591	−1.290
				(1.404)	(1.443)
Fraction 50-59				1.047	0.766
				(1.095)	(1.100)
Peak	14.405	14.413	14.075	13.23	13.286
	(0.321)	(0.219)	(0.156)	(0.486)	(0.310)
Observations	1,632	1,632	1,632	1,632	1,632
R^2	0.091	0.118	0.123	0.098	0.129
Adjusted R^2	0.084	0.110	0.114	0.089	0.118
Note:					*p<0.1; **p<0.05; ***p<0.01

Note: This table shows a hump-shaped relationship between age diversity and GDP per capita. The specification regresses first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.8: Predicted age diversity and GDP per capita (first-differenced)

	Dependent variable:				
	Log GDP per capita				
	(1)	(2)	(3)	(4)	(5)
Pred. age div.	0.314**	0.381**	0.544**	0.365**	0.839***
	(0.147)	(0.166)	(0.216)	(0.160)	(0.192)
Pred. age div. sq.	−0.014**	−0.016**	−0.022**	−0.015**	−0.032***
	(0.007)	(0.007)	(0.009)	(0.007)	(0.008)
Pred. young-age dep.		−0.239*	−0.045		0.198***
		(0.133)	(0.107)		(0.066)
Pred. old-age dep.		0.313	−0.470		−1.833***
		(0.201)	(0.498)		(0.398)
Pred. mean age			0.050*		0.132***
			(0.028)		(0.025)
Pred. 20-29				−0.455	2.351*
				(1.293)	(1.294)
Pred. 30-39				0.309	2.341
				(1.554)	(1.526)
Pred. 40-49				0.814	1.716
				(1.596)	(1.556)
Pred. 50-59				0.224	−0.159
				(1.462)	(1.414)
Peak	11.156	11.757	12.262	12.159	13.313
	(0.076)	(0.321)	(0.297)	(1.79)	(0.6)
Observations	802	802	802	802	802
R^2	0.090	0.097	0.107	0.095	0.121
Adjusted R^2	0.082	0.087	0.096	0.082	0.106
Note:					*p<0.1; **p<0.05; ***p<0.01

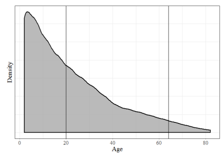
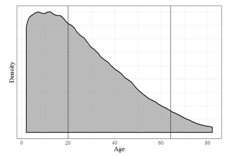
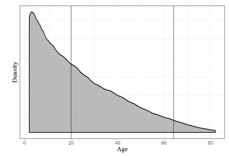
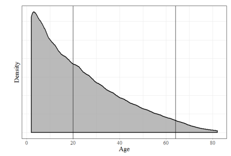
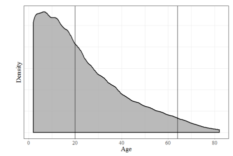
Note: This table shows a hump-shaped relationship between predicted age diversity and GDP per capita. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specification regresses first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.9: Instrumented age diversity and GDP per capita (first-differenced)

	Dependent variable:				
	Log GDP per capita				
	(1)	(2)	(3)	(4)	(5)
Age diversity	4.176	4.719*	1.031	3.344***	2.699***
	(2.832)	(2.854)	(2.557)	(0.937)	(0.831)
Age diversity sq.	−0.163	−0.184*	−0.044	−0.127***	−0.107***
	(0.108)	(0.108)	(0.096)	(0.035)	(0.031)
Young-age dependency		−0.670***	−0.062		−0.068
		(0.167)	(0.357)		(0.500)
Old-age dependency		3.396***	−0.867		−2.522
		(1.218)	(1.530)		(1.914)
Mean age			0.085**		0.087
			(0.040)		(0.063)
Fraction 20-29				−1.556	−3.108
				(3.233)	(5.897)
Fraction 30-39				−1.072	−3.981
				(4.355)	(7.369)
Fraction 40-49				0.044	−2.840
				(3.585)	(5.901)
Fraction 50-59				−0.592	−3.178
				(2.195)	(3.694)
Peak	12.795	12.828	11.657	13.184	12.570
	(0.088)	(0.072)	(14.473)	(0.448)	(1.082)
First-stage F	139.4	135.0	73.9	35.9	28.7
Observations	669	669	669	669	669
R^2	0.080	0.106	0.097	0.090	0.116
Adjusted R^2	0.071	0.095	0.085	0.076	0.098
Note:	*p<0.1; **p<0.05; ***p<0.01				

Note: This table shows a hump-shaped relationship between age diversity and GDP per capita with age diversity instrumented by predicted age diversity. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specification regresses first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.10: Examples of population structures around the optimal level of diversity

	 COM 1990	 ZAF 2000	 CRI 1950	 ETH 1970	 LBY 1990
MAD	13.312	13.314	13.312	13.311	13.314
SD	11.954	11.839	11.854	11.884	11.940
Gini	0.187	0.183	0.184	0.185	0.187
HHI	0.860	0.868	0.866	0.864	0.860
GE	0.057	0.056	0.056	0.056	0.057
Atkinson	9.996	9.772	9.866	9.910	10.042

Note: This table shows the population structures of countries whose diversity (measured by MAD) is closest to the optimal level (13.313). Each column is for a different country/year pair. The countries are “COM” for the Comoros, “ZAF” for South Africa, “CRI” for Costa Rica, “ETH” for Ethiopia, and “LBY” for Libya. The graphs show the population’s age distribution. The two vertical lines in each graph are at age 20 and 64, and they denote the bounds of the working-age population.

Table D.11: The channel behind the hump-shape (actual age diversity)

	Dependent variable:					
	TFP		Capital per capita		Human capital	
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	1.374^{***}	1.535^{***}	-0.714	-0.790	-0.345^{**}	-0.327^{*}
	(0.387)	(0.362)	(2.073)	(2.353)	(0.155)	(0.170)
Age diversity sq.	-0.049^{***}	-0.058^{***}	0.035	-0.010	0.013^{**}	0.013^{**}
	(0.014)	(0.013)	(0.076)	(0.088)	(0.006)	(0.006)
Young-age dependency		0.391^{**}		5.089^{**}		-0.039
		(0.181)		(2.440)		(0.092)
Old-age dependency		-1.092^{*}		3.666		-0.407
		(0.585)		(12.275)		(0.284)
Mean age		0.054^{***}		0.714^{*}		0.033^{***}
		(0.020)		(0.420)		(0.011)
Fraction 20-29		-0.065		-14.846		1.008^{**}
		(0.942)		(10.090)		(0.415)
Fraction 30-39		-1.006		-27.743^{***}		0.942^{**}
		(1.248)		(9.541)		(0.436)
Fraction 40-49		-1.711		-25.533^{***}		0.998^{***}
		(1.319)		(6.687)		(0.384)
Fraction 50-59		-0.241		-15.629^{***}		0.634^{*}
		(0.725)		(4.631)		(0.344)
Peak	14.124	13.3	10.308	-40.694	13.345	13.038
	(0.055)	(0.153)	(57.133)	(240913.786)	(0.177)	(0.624)
Observations	1,068	1,068	1,629	1,629	1,421	1,421
R^2	0.129	0.143	0.136	0.200	0.102	0.175
Adjusted R^2	0.119	0.126	0.129	0.190	0.094	0.163

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the hump-shaped relationship between age diversity and GDP per capita is driven by the productivity (TFP) channel. The specifications regress first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.12: The channel behind the hump-shape (predicted age diversity)

	Dependent variable:					
	TFP		Capital per capita		Human capital	
	(1)	(2)	(3)	(4)	(5)	(6)
Pred. age div.	0.120	0.703***	−2.220	−0.831	−0.104	0.008
	(0.106)	(0.229)	(2.132)	(1.530)	(0.088)	(0.079)
Pred. age div. sq.	−0.005	−0.031***	0.084	−0.010	0.004	0.001
	(0.004)	(0.009)	(0.092)	(0.070)	(0.003)	(0.003)
Pred. young-age dep.		0.244***		0.352		−0.162***
		(0.069)		(0.916)		(0.022)
Pred. old-age dep.		−1.727***		6.205		−0.106
		(0.330)		(5.042)		(0.174)
Pred. mean age		0.077***		0.077		0.006
		(0.017)		(0.189)		(0.011)
Pred. 20-29		−1.450*		−54.509***		−0.199
		(0.812)		(12.308)		(0.396)
Pred. 30-39		−2.973**		−65.039***		0.083
		(1.356)		(13.127)		(0.461)
Pred. 40-49		−2.881**		−53.184***		0.619
		(1.163)		(12.073)		(0.393)
Pred. 50-59		−1.956**		−55.138***		−0.142
		(0.862)		(11.280)		(0.388)
Peak	12.431	11.477	13.182	-41.839	12.812	-6.307
	(1.313)	(0.34)	(5.514)	(136173.584)	(1.504)	(8434.188)
Observations	594	594	802	802	742	742
R^2	0.101	0.151	0.139	0.213	0.021	0.132
Adjusted R^2	0.090	0.130	0.131	0.199	0.012	0.115

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the hump-shaped relationship between predicted age diversity and GDP per capita is driven by the productivity (TFP) channel. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specifications regress first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.13: The channel behind the hump-shape (instrumented age diversity)

	Dependent variable:					
	TFP		Capital per worker		Human capital	
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	1.634*	0.580	−14.141	−3.086	−0.050	−0.050
	(0.868)	(0.461)	(32.606)	(8.725)	(0.629)	(0.440)
Age diversity sq.	−0.062*	−0.021	0.513	−0.104	0.001	−0.005
	(0.033)	(0.017)	(1.231)	(0.387)	(0.023)	(0.014)
Young-age dependency		0.004		3.353		−0.210
		(0.259)		(7.969)		(0.197)
Old-age dependency		0.595		−11.190		−1.278*
		(1.159)		(24.220)		(0.750)
Mean age		−0.008		0.719		0.020
		(0.029)		(1.037)		(0.022)
Fraction 20-29		0.148		−142.056		−3.578
		(2.677)		(100.001)		(4.289)
Fraction 30-39		0.861		−200.367		−4.752
		(3.619)		(122.190)		(5.525)
Fraction 40-49		0.709		−165.458*		−3.702
		(2.814)		(98.204)		(4.352)
Fraction 50-59		1.331		−120.203**		−2.380
		(1.716)		(59.875)		(2.756)
Peak	13.082	14.066	13.775	-14.846	34.701	-5.531
	(0.059)	(9.102)	(1.963)	(9213.272)	(476038.997)	(4222.731)
First-stage F	139.4	28.7	139.4	28.7	139.4	28.7
Observations	495	495	669	669	619	619
R^2	0.033	0.047	0.115	0.231	0.012	0.108
Adjusted R^2	0.021	0.022	0.107	0.216	0.003	0.089
Note:						*p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the hump-shaped relationship between age diversity (instrumented by predicted age diversity) and GDP per capita is driven by the productivity (TFP) channel. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specifications regress first differences on first differences. All columns include year fixed effects. Robust standard errors are clustered at the country-level.

Table D.14: Alternative measures of age diversity (actual)

	Dependent variable:				
	Log GDP per capita				
	SD	Gini	HHI 5	HHI 10	HHI 15
	(1)	(2)	(3)	(4)	(5)
Age diversity	1.787***	1.570***	38.802	−6.872	−16.146
	(0.678)	(0.505)	(195.917)	(43.252)	(16.906)
Age diversity sq.	−0.077***	−0.043***	−18.463	8.060	14.769
	(0.028)	(0.014)	(114.798)	(29.918)	(14.449)
Peak	11.678	18.173	1.051	0.426	0.547
	(0.185)	(0.182)	(1.524)	(1.225)	(0.004)
Median diversity	12.096	18.358	0.870	0.744	0.618
All controls	Yes	Yes	Yes	Yes	Yes
Observations	1,632	1,632	1,632	1,632	1,632
R^2	0.130	0.131	0.124	0.125	0.123
Adjusted R^2	0.119	0.120	0.113	0.114	0.112
Note:					*p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the hump-shaped relationship between age diversity and GDP per capita is robust to alternative measures of age diversity. "HHI X" means that the HHI was calculated by splitting the working-age population into X-year age groups. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.15: Generalised entropy and Atkinson indices (actual)

	Dependent variable:					
	Log GDP per capita					
	Generalised entropy			Atkinson		
	$\alpha = -1.4$	$\alpha = 0.0$	$\alpha = 1.0$	$\varepsilon = 2.0$	$\varepsilon = 3.0$	$\varepsilon = 4.0$
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	1.123^{**}	1.442^{***}	1.350^{***}	0.860^{**}	0.605^{**}	0.476^{**}
	(0.480)	(0.542)	(0.494)	(0.342)	(0.258)	(0.214)
Age diversity sq.	-0.100^{**}	-0.139^{***}	-0.131^{***}	-0.044^{**}	-0.021^{**}	-0.013^{**}
	(0.042)	(0.051)	(0.047)	(0.017)	(0.009)	(0.006)
Peak	5.632	5.196	5.133	9.831	14.090	17.693
	(0.059)	(0.047)	(0.063)	(0.135)	(0.280)	(0.491)
Median diversity	5.700	5.275	5.213	9.972	14.271	17.896
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,632	1,632	1,632	1,632	1,632	1,632
R^2	0.128	0.129	0.129	0.129	0.129	0.128
Adjusted R^2	0.118	0.119	0.118	0.118	0.118	0.117
Note:				*p<0.1; **p<0.05; ***p<0.01		

Note: This table shows that the hump-shaped relationship between age diversity and GDP per capita is robust to alternative measures of age diversity. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.16: Alternative measures of age diversity (predicted)

	Dependent variable:				
	Log GDP per capita				
	SD	Gini	HHI 5	HHI 10	HHI 15
	(1)	(2)	(3)	(4)	(5)
Pred. age div.	1.479***	0.561***	61.273***	30.921**	6.453
	(0.354)	(0.128)	(22.430)	(12.795)	(11.381)
Pred. age div. sq.	−0.062***	−0.017***	−38.311***	−21.189**	−3.486
	(0.016)	(0.005)	(14.691)	(9.593)	(9.731)
Peak	11.997	16.878	0.800	0.730	0.926
	(0.485)	(2.014)	(0.000)	(0.001)	(0.921)
Median diversity	18.352	11.988	0.868	0.740	0.612
All controls	Yes	Yes	Yes	Yes	Yes
Observations	802	802	802	802	802
R^2	0.120	0.122	0.124	0.120	0.113
Adjusted R^2	0.104	0.106	0.108	0.104	0.098
Note:	*p<0.1; **p<0.05; ***p<0.01				

Note: This table shows that the hump-shaped relationship between predicted age diversity and GDP per capita is robust to alternative measures of age diversity. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. “HHI X” means that the HHI was calculated by splitting the working-age population into X-year age groups. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.17: Generalised entropy and Atkinson indices (predicted)

	Dependent variable:					
	Log GDP per capita					
	Generalised entropy			Atkinson		
	$\alpha = -1.4$	$\alpha = 0.0$	$\alpha = 1.0$	$\varepsilon = 2.0$	$\varepsilon = 3.0$	$\varepsilon = 4.0$
	(1)	(2)	(3)	(4)	(5)	(6)
Pred. age div.	0.986***	1.000***	0.918***	0.657***	0.555***	0.524***
	(0.291)	(0.274)	(0.253)	(0.174)	(0.155)	(0.164)
Pred. age div. sq.	−0.088***	−0.098***	−0.093***	−0.034***	−0.020***	−0.015***
	(0.025)	(0.028)	(0.028)	(0.009)	(0.005)	(0.004)
Peak	5.591	5.085	4.953	9.707	14.006	17.653
	(0.193)	(0.263)	(0.401)	(0.567)	(0.810)	(1.079)
Median diversity	5.636	5.255	5.218	9.892	14.113	17.681
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	802	802	802	802	802	802
R^2	0.116	0.117	0.118	0.118	0.117	0.116
Adjusted R^2	0.101	0.102	0.102	0.102	0.101	0.100

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the hump-shaped relationship between predicted age diversity and GDP per capita is robust to alternative measures of age diversity. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.18: Alternative measures of age diversity (instrumented)

	Dependent variable:				
	Log GDP per capita				
	SD	Gini	HHI 5	HHI 10	HHI 15
	(1)	(2)	(3)	(4)	(5)
Age diversity	3.703***	1.618**	−2.068	−20.515	−18.920
	(1.115)	(0.741)	(2.273)	(52.644)	(24.742)
Age diversity sq.	−0.170***	−0.049***	0.012	17.404	19.741
	(0.041)	(0.019)	(0.013)	(35.297)	(20.230)
Peak	10.879	16.417	84.774	0.589	0.479
	(1.630)	(3.992)	(15.792)	(0.113)	(0.026)
Median diversity	12.096	18.358	86.991	0.744	0.618
First-stage F	24.3	32.4	13.9	26.9	70.9
Observations	669	669	669	669	669
R^2	0.119	0.112	0.103	0.103	0.105
Adjusted R^2	0.101	0.095	0.086	0.085	0.087
Note:	*p<0.1; **p<0.05; ***p<0.01				

Note: This table shows that the hump-shaped relationship between age diversity (instrumented by predicted age diversity) and GDP per capita is robust to alternative measures of age diversity. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. "HHI X" means that the HHI was calculated by splitting the working-age population into X-year age groups. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.19: Generalised entropy and Atkinson indices (instrumented)

	Dependent variable:					
	Log GDP per capita					
	-1.4	0.0	1.0	2.0	3.0	4.0
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	2.145^{***}	1.692^{*}	0.968	1.461^{***}	1.284^{***}	1.044^{***}
	(0.771)	(0.975)	(1.022)	(0.540)	(0.377)	(0.306)
Age diversity sq.	-0.226^{***}	-0.210^{***}	-0.153^{**}	-0.087^{***}	-0.052^{***}	-0.034^{***}
	(0.050)	(0.067)	(0.069)	(0.020)	(0.011)	(0.008)
Peak	4.746	4.021	3.157	8.401	12.353	15.489
	(0.858)	(1.701)	(4.812)	(2.495)	(3.431)	(5.965)
Median diversity	5.700	5.275	5.213	9.972	14.271	17.896
First-stage F	35.1	35.0	33.7	33.9	32.2	30.0
Observations	669	669	669	669	669	669
R^2	0.121	0.116	0.113	0.120	0.121	0.119
Adjusted R^2	0.103	0.099	0.095	0.102	0.104	0.101
Note:	*p<0.1; **p<0.05; ***p<0.01					

Note: This table shows that the hump-shaped relationship between age diversity (instrumented by predicted age diversity) and GDP per capita is robust to alternative measures of age diversity. Predicted age diversity is calculated from UN population forecasts from 1982 as opposed to actual population data. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.20: Age diversity and GDP per capita in European regions

	Dependent variable:					
	Log GDP per capita					
	MAD	SD	Gini	HHI 5	GE $\alpha = -0.7$	Atkinson $\varepsilon = 1.7$
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	2.276**	3.002***	1.270*	10.636	2.252***	1.644***
	(0.914)	(1.140)	(0.675)	(8.627)	(0.698)	(0.508)
Age diversity sq.	−0.082**	−0.123***	−0.039*	−0.060	−0.221***	−0.102***
	(0.033)	(0.047)	(0.020)	(0.049)	(0.069)	(0.032)
Peak	13.913	12.205	16.261	88.672	5.093	8.058
	(0.133)	(0.093)	(0.496)	(0.407)	(0.035)	(0.076)
Median diversity	14.237	12.426	17.099	88.676	5.161	8.163
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,316	4,316	4,316	4,316	4,316	4,316
R^2	0.983	0.983	0.983	0.983	0.983	0.983
Adjusted R^2	0.981	0.982	0.981	0.981	0.982	0.982
Note:						*p<0.1; **p<0.05; ***p<0.01

Note: This table shows the hump-shaped relationship between age diversity and GDP per capita for European NUTS2 regions for six different measures of age diversity. The specifications regress levels on levels. All columns include year, region, and country fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the region-level.

Table D.21: Age diversity and GDP per capita in European countries

	Dependent variable:					
	Log GDP per capita					
	MAD	SD	Gini	HHI 5	GE $\alpha = -0.7$	Atkinson $\varepsilon = 1.7$
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	10.328***	12.962***	9.259***	−205.391	9.919***	7.123***
	(3.970)	(4.901)	(3.085)	(246.622)	(3.002)	(2.191)
Age diversity sq.	−0.350**	−0.507**	−0.263***	1.163	−0.888***	−0.405***
	(0.141)	(0.199)	(0.090)	(1.392)	(0.283)	(0.131)
Peak	14.760	12.792	17.596	88.304	5.587	8.794
	(0.327)	(0.203)	(0.269)	(0.158)	(0.064)	(0.146)
Median diversity	14.349	12.513	17.347	88.728	5.302	8.369
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	469	469	469	469	469	469
R^2	0.983	0.983	0.985	0.983	0.985	0.985
Adjusted R^2	0.981	0.981	0.983	0.981	0.983	0.983
Note:						*p<0.1; **p<0.05; ***p<0.01

Note: This table shows the hump-shaped relationship between age diversity and GDP per capita for European countries for six different measures of age diversity. The specifications regress levels on levels. All columns include year, region, and country fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.22: Robustness to flexible mean age control (country-level)

	Dependent variable:				
	Log GDP per capita				
	MAD	SD	Gini	GE $\alpha = -0.7$	Atkinson $\varepsilon = 1.7$
	(1)	(2)	(3)	(4)	(5)
Age diversity	2.700^{***}	3.945^{***}	2.030^{*}	1.898^{*}	1.174^{**}
	(0.867)	(1.152)	(1.066)	(0.987)	(0.502)
Age diversity sq.	-0.106^{***}	-0.173^{***}	-0.058^{*}	-0.184^{**}	-0.045^{**}
	(0.032)	(0.048)	(0.029)	(0.085)	(0.017)
Peak	12.722	11.403	17.642	5.157	13.133
	(0.285)	(0.177)	(0.581)	(0.193)	(0.652)
Median diversity	13.630	12.096	18.358	5.700	14.271
Observations	1,805	1,805	1,805	1,805	1,805
R^2	0.931	0.931	0.930	0.930	0.931
Adjusted R^2	0.922	0.923	0.921	0.922	0.922
Note:	*p<0.1; **p<0.05; ***p<0.01				

Note: This table shows that the hump-shaped relationship between age diversity and GDP per capita is robust to adding flexible control for the mean age of the working age population in the country-level analysis. The specifications regress levels on levels. All columns include year, region, and country fixed effects, and controls for the old- and young-age dependency ratios, mean age (up to a fifth-degree polynomial), and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level.

Table D.23: Robustness to flexible mean age control (European regions)

	Dependent variable:				
	Log GDP per capita				
	MAD	SD	Gini	GE $\alpha = -0.7$	Atkinson $\varepsilon = 1.7$
	(1)	(2)	(3)	(4)	(5)
Age diversity	2.560**	3.288**	1.207*	2.229***	1.639***
	(1.072)	(1.312)	(0.682)	(0.731)	(0.530)
Age diversity sq.	$-$0.095**	$-$0.139***	$-$0.040*	$-$0.226***	$-$0.105***
	(0.038)	(0.054)	(0.020)	(0.071)	(0.033)
Peak	13.436	11.867	15.143	4.928	7.817
	(0.172)	(0.101)	(1.176)	(0.044)	(0.094)
Median diversity	14.237	12.426	17.099	5.161	8.163
Observations	4,316	4,316	4,316	4,316	4,316
R^2	0.983	0.983	0.983	0.983	0.983
Adjusted R^2	0.982	0.982	0.982	0.982	0.982
Note:			*p<0.1; **p<0.05; ***p<0.01		

Note: This table shows that the hump-shaped relationship between age diversity and GDP per capita is robust to adding flexible control for the mean age of the working age population in the European regional analysis. The specifications regress levels on levels. All columns include year, region, and country fixed effects, and controls for the old- and young-age dependency ratios, mean age (up to a fifth-degree polynomial), and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the region-level.

Table D.24: Optimal level of age diversity and returns to education

	Dependent variable:					
	Log GDP per capita					
	Returns to edu.			Growth in returns		
	Actual	Predicted	Instrumented	Actual	Predicted	Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)
Age diversity	0.199	0.331	1.223	0.168	1.084*	0.442
	(1.700)	(1.351)	(1.738)	(0.783)	(0.628)	(1.708)
Age div. sq.	−0.005	−0.015	−0.078	−0.005	−0.044*	−0.063
	(0.061)	(0.052)	(0.057)	(0.027)	(0.024)	(0.062)
Age div. $\times r_{ct}$	0.002	0.004	−0.006	1.604***	−129.618	−6.750
	(0.031)	(0.021)	(0.010)	(0.216)	(86.393)	(10.684)
Age div. sq. $\times r_{ct}$	−0.0001	−0.0002	0.0002	−0.062***	4.783	0.322
	(0.001)	(0.001)	(0.0004)	(0.009)	(3.222)	(0.478)
r_{ct}	−0.004	−0.020	0.049	−10.308***	878.047	33.494
	(0.220)	(0.138)	(0.078)	(1.290)	(579.092)	(58.125)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	-	-	32.292	-	-	33.143
Observations	1,133	504	504	1,081	504	504
R^2	0.119	0.120	0.122	0.119	0.123	0.123
Adjusted R^2	0.100	0.091	0.093	0.101	0.094	0.094

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the optimal level of age diversity depends on returns to education and in the growth rate of returns to education. The specifications regress first differences on first differences. All columns include year fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the country-level. The columns labelled “Actual” use actual age diversity, “Predicted” columns use age diversity calculated from UN population forecasts from 1982, “Instrumented” columns instrument actual age diversity with predicted age diversity.

Table D.25: Variation of optimal age diversity with returns to education

		Quantile of r_{ct}										
		0	10	20	30	40	50	60	70	80	90	100
Returns to education	Lin. coef.	0.199	0.204	0.210	0.217	0.225	0.235	0.248	0.266	0.289	0.318	0.361
	Quad. coef.	-0.005	-0.005	-0.005	-0.006	-0.006	-0.007	-0.007	-0.008	-0.009	-0.011	-0.013
	Peak	21.038	20.522	19.941	19.294	18.641	17.929	17.222	16.397	15.563	14.802	13.966
Growth of returns	Lin. coef.	0.152	0.166	0.168	0.170	0.171	0.172	0.174	0.176	0.180	0.191	24.217
	Quad coef.	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.006	-0.006	-0.941
	Peak	16.782	16.359	16.294	16.255	16.223	16.191	16.152	16.097	16.007	15.784	12.872
High-tech. employment	Lin. coef.	4.804	4.141	3.754	3.477	3.201	2.980	2.704	2.372	2.040	1.322	-2.215
	Quad. coef.	-0.171	-0.147	-0.133	-0.123	-0.114	-0.106	-0.096	-0.084	-0.072	-0.046	0.081
	Peak	14.040	14.057	14.070	14.082	14.095	14.107	14.125	14.153	14.190	14.335	13.700
Manufact. employment	Lin. coef.	5.129	3.672	3.187	2.764	2.384	2.025	1.603	1.160	0.463	-0.318	-3.317
	Quad. coef.	-0.180	-0.129	-0.112	-0.097	-0.084	-0.071	-0.056	-0.041	-0.016	0.011	0.117
	Peak	14.218	14.224	14.227	14.230	14.235	14.240	14.250	14.268	14.367	13.972	14.181
Low-tech. employment	Lin. coef.	5.731	4.280	3.777	3.386	2.996	2.661	2.214	1.711	1.041	-0.075	-7.278
	Quad. coef.	-0.204	-0.152	-0.134	-0.120	-0.107	-0.095	-0.079	-0.061	-0.037	0.003	0.261
	Peak	14.023	14.037	14.044	14.052	14.061	14.071	14.089	14.121	14.212	11.428	13.950
Agricult. employment	Lin. coef.	0.120	0.691	1.025	1.311	1.692	2.121	2.645	3.597	5.122	9.791	31.133
	Quad. coef.	-0.004	-0.024	-0.035	-0.045	-0.058	-0.073	-0.091	-0.124	-0.176	-0.336	-1.068
	Peak	13.519	14.387	14.450	14.479	14.502	14.518	14.531	14.545	14.556	14.569	14.578

Note: This table shows how the linear and quadratic coefficients of age diversity as well as the optimal level of age diversity depend on returns to education and its growth rate (based on the country-level regressions with actual age diversity), and employment in different sectors (based on the EU regressions). Each column corresponds to the coefficients and optimal diversity at a different quantile of returns or employment going from the sample minimum (0th quantile) to the sample maximum (100th quantile). The linear coefficients are broadly positive, while the quadratic ones negative indicating that the relationship is hump-shaped. The peaks are decreasing in returns to education and their growth rate in accordance with the model. The peaks also increase more when an economy is more reliant on low-skill sectors such as agriculture indicating that high-skill sectors are more amenable to lower levels of diversity.

Table D.26: Optimal level of age diversity in different sectors (European regions)

	Dependent variable:			
	Log GDP per capita			
	High-tech.	Manufacturing	Low-tech.	Agriculture
	(1)	(2)	(3)	(4)
Age diversity	4.859^{*}	5.319^{*}	6.011^{***}	0.072
	(2.547)	(2.935)	(2.321)	(1.395)
Age diversity sq.	-0.173^{*}	-0.187^{*}	-0.214^{***}	-0.003
	(0.090)	(0.104)	(0.082)	(0.049)
Employment	3.869	1.503	3.892^{*}	-3.497^{**}
	(3.509)	(1.171)	(2.297)	(1.490)
Age div. $\times$ Emp.	-0.553	-0.211	-0.558^{*}	0.476^{**}
	(0.501)	(0.166)	(0.324)	(0.206)
Age div. sq. $\times$ Emp.	0.020	0.007	0.020^{*}	-0.016^{**}
	(0.018)	(0.006)	(0.011)	(0.007)
Peak, low emp.	14.040	14.218	14.023	13.519
Peak, high emp.	13.700	14.181	13.950	14.578
All controls	Yes	Yes	Yes	Yes
Observations	3,773	4,031	3,999	3,847
R^2	0.984	0.982	0.982	0.984
Adjusted R^2	0.982	0.980	0.980	0.982

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows that the optimal level of age diversity is negatively associated with employment in skill-intensive sectors in European regions. Two peaks are shown, one for the sample minimum of employment ("low emp."), and one for the sample maximum ("high emp."). The specifications regress levels on levels. All columns include year, region, and country fixed effects, and controls for the old- and young-age dependency ratios, mean age, and the share of 20-29, 30-39, 40-49, and 50-59-year-olds in the working-age population. Robust standard errors are clustered at the region-level.

E Data description

E.1 Dependent variables

Log GDP per capita (countries): This is the main dependent variable. It refers to expenditure-side real GDP divided by population, both from the Penn World Table 9.0 (Feenstra et al. (2015)). The data covers years 1950-2014.

Log GDP per capita (European regions): real GDP per capita in EUR from Eurostat. The data covers the years 2000-2015.

Total factor productivity: real total factor productivity from the Penn World Table 9.0 (Feenstra et al. (2015)). The data covers years 1950-2014.

Capital per capita: real capital stock divided by population from the Penn World Table 9.0 (Feenstra et al. (2015)). The data covers years 1950-2014.

Human capital: human capital index from the Penn World Table 9.0 (Feenstra et al. (2015)). This is based on combining average years of schooling data from Lee and Lee (2016) and Cohen and Leker (2014). The data covers years 1950-2014. For more information, see Penn World Table (2014).

E.2 Explanatory variables

Age diversity (countries): This is the main explanatory variable. The mean absolute difference, standard deviation, Gini coefficient, Herfindahl-Hirschman Index, Generalised Entropy index, or Atkinson index (as indicated) of the working-age (20-64) population. This is based on population data in five-year age groups from the United Nations (2017) encompassing the 1950-2015 period.

Young-age dependency (countries): number of 0-19-year-olds divided by the number of working-age (20-64) people. This is based on population data from the United Nations (2017) encompassing the 1950-2015 period.

Old-age dependency (countries): number of people over age 64 divided by the number of working-age (20-64) people. This is based on population data from the United Nations (2017) encompassing the 1950-2015 period.

Mean age (countries): average age of the working-age (20-64) population. This is based on population data from the United Nations (2017) encompassing the 1950-2015 period.

Fraction X-Y (countries): number of people in age group X-Y divided by the working-age (20-64) population. This is based on population data from the United Nations (2017) encompassing the 1950-2015 period.

Predicted age diversity: the mean absolute difference, standard deviation, Gini coefficient, Herfindahl-Hirschman Index, Generalised Entropy index, or Atkinson index (as indicated) of the working-age (20-64) population. This is based on projected population data in five-year age groups from the United Nations (1985) encompassing the 1985-2015 period.

Predicted young-age dependency: number of 0-19-year-olds divided by the number of working-age (20-64) people. This is based on projected population data from the United Nations (1985) encompassing the 1985-2015 period.

Predicted old-age dependency: number of people over age 64 divided by the number of working-age (20-64) people. This is based on projected population data from the United Nations (1985) encompassing the 1985-2015 period.

Predicted mean age: average age of the working-age (20-64) population. This is based on projected population data from the United Nations (1985) encompassing the 1985-2015 period.

Predicted fraction X-Y: number of people in age group X-Y divided by the working-age (20-64) population. This is based on projected population data from the United Nations (1985) encompassing the 1985-2015 period.

Age diversity (European regions): the mean absolute difference, standard deviation, Gini coefficient, Herfindahl-Hirschman Index, Generalised Entropy index, or Atkinson index (as indicated) of the working-age (20-64) population. This is based on population data in five-year age groups from the Eurostat encompassing the 1990-2016 period.

Young-age dependency (European regions): number of 0-19-year-olds divided by the number of working-age (20-64) people. This is based on population data in five-year age groups from the Eurostat encompassing the 1990-2016 period.

Old-age dependency (European regions): number of people over age 64 divided by the number of working-age (20-64) people. This is based on population data in five-year age groups from the Eurostat encompassing the 1990-2016 period.

Mean age (European regions): average age of the working-age (20-64) population. This is based on population data in five-year age groups from the Eurostat encompassing the 1990-2016 period.

Fraction X-Y (European regions): number of people in age group X-Y divided by the working-age (20-64) population. This is based on population data in five-year age groups from the Eurostat encompassing the 1990-2016 period.

Returns to education: proxied by the share of people in the population with more than primary educational attainment. This is constructed at five-year intervals from Lee and Lee (2016) spanning 1950-2010.

Growth in returns to human capital: share of 20-24-year-olds with tertiary education divided by the share of 20-24-year-olds with tertiary education five years earlier. This is constructed at five-year intervals from Lee and Lee (2016) following Ang and Madsen (2015) spanning 1950-2010.

Employment in a sector: number of people employed in a given sector divided by the number of all employed people. This is available for European regions from the Eurostat for 1999-2017.