**WORKING PAPERS**

# Economics breeds culture

Patricio Garcia-Mínguez

Ausias Ribo Argemi

Universitat de Barcelona

# Economics breeds culture

**Abstract**: Several recent papers document the influence and long lasting effects of technology on preferences. Simultaneously, cultural factors are often invoked to explain heterogeneity in preferences. These two ideas suggest that culture determines the short run equilibrium values of economic variables, but, in the long run, culture changes in response to the underlying economic fundamentals. We build a model in which preferences are endogenous and the diversity in preferences (the "cultural" diversity) is explained by the variation in the relevant economic fundamentals. This can help explain observed differences in labor market attachment among groups defined e.g., by citizenship, ethnicity or gender.

Patricio Garcia-Mínguez
Facultat d'Economia i Empresa
Universitat de Barcelona

Ausias Ribo Argemi
Facultat d'Economia i Empresa
Universitat de Barcelona

.

# 1   Introduction

Why fewer women than men do market work? Why Koreans or Americans work longer hours than Europeans? Ask a sociologist and she will likely point to cultural heterogeneity as the main explanation of these differences. While recognizing the importance of culture, economists tend to understand better explanations based on opportunity costs or incentives and try to avoid those involving differences in preferences (of which we tend to be suspicious at least since Stigler and Becker, 1977). One problem with differences in preferences is that, although tastes are clearly heterogeneous across individuals, it is hard to understand why a specific preference should be more frequent among individuals living in a particular country or belonging to a particular group. However, Fernández (2010) surveys a number of papers that provide ample cross-sectional, historical and experimental evidence that this is indeed the case.

Our understanding is that culture shapes individual preferences (and, therefore, it should not be surprising to find group or geographical clustering of preferences), but this begets the question of what gave rise to that particular culture in the first place. We propose an explanation in which it is economic factors (more specifically, technology) that generate a culture that, in turn, shapes the individual preferences. Culture, thus, determines through preferences the short run equilibrium values of economic variables. In the long run, however, it is culture what changes in response to the underlying economic fundamentals.

Previous efforts to understand the origin of preferences have stressed the central role that exposure and experience play in the development of tastes.[1] Of course, the notion that preferences are shaped by habit and custom is not new in Economics, as it can be traced back at least to Marshall (1920). In this origin-of-preferences literature, two strands can be identified, that of "external" and that of "internal habits." We use "external habits" to denote the idea that it is unintended exposure to a given good or service what helps the individual to develop a taste (or aversion) to it. It is conventional to cite Duesenberry (1949) as the precursor of this habit formation theory that adopted its canonical form under Constantinides (1990) and Abel (1990), and was latter extended to allow for intergenerational interactions by de la Croix (1996). In a similar vein, "internal habits" refers to those cases in which it is the current choices of the individual what shape her future preferences, because learning by consuming takes place. This literature originated in Stigler and Becker (1977) and was later developed by Becker and Murphy (1988).

As far as what the ultimate origin of preferences is, this "habits" theory is wanting because the preferences of an individual end up being explained, at least partially, by the preferences of other people, and this naturally elicits the question of what is the origin of these other people's

---

[1]See Bowles (1998) for evidence taken from other disciplines on the relevance of exposure in the development of tastes.

preferences. We believe that searching for the origin of preferences in preferences themselves is, to some extent, tautological and we need to look at other fundamentals if we want to understand their genesis. We share the view of Bowles (1998) that some basic preferences might be innate, possibly genetically transmitted, while the rest would be learned. Those innate would be very general in nature, such as the preference for sweet tastes or the dislike of pain, and probably are the result of the gene selection process in the evolution of the species.[2] The real problem lies in understanding why a particular culture arises in which we learn to appreciate specific qualities of the goods.

We posit that this is governed by endowments and technology, or, more precisely, by their interaction. For example, if we look at fish and seafood consumption per capita, we observe that it is highest in tiny Pacific island nations and lowest in some landlocked countries. Using country data, the correlation coefficient between fish and seafood consumption per year per capita and the ratio of coastline length to total land area is of 0,66 with a t-statistic of 8,64. Even excluding Maldives, whose exceptionally large per capita consumption may be driven by tourism, the correlation coefficient remains substantial with a value of 0,47 (6,09).[3] A plausible rationale is that, where fish was readily available, individuals developed a taste for fish, something that did not happen where it was not, and this "cultural preference" is driving consumption decisions today, when improved transportation methods have increased dramatically fish accessibility everywhere. If our intuition is correct, it would help understand why we observe that people living in a given country or belonging to a given group tend to have similar preferences.

It should be noted that this "availability" could be the result of either endowments or technology. In our earlier example, fish could be easily accesible in a particular area because fish stocks were extremely abundant and even a rudimentary fishing technology yields a large production. In this case, fish availability for consumption would be due to the endowment. Alternatively, a more refined fishing technology could yield the same production from far less plentiful stocks. In this case it would be due to technology. In either case, the product of labor could be the same and, to the extent that this is the case, the two alternatives would be indistinguishable. Hence, it is the interaction of endowments and technology what, in our view, shapes preferences.

As far as technology is concerned, there is a budding empirical literature dealing with the technology-to-preferences line of causation. Alesina et al. (2010) provide a nice example: they uncover a link between historic plough use and current self expressed attitudes about the role of women in society. Their argument is that plough use increased the productivity of males

---

[2]Some tastes, e.g., bitter taste, are asociated with natural poisons, e.g., hemlock, so it is natural that avoiding those tastes became favored by natural selection. However, the learning process may well trump that innate taste and help develop an apreciation for tonic water or, at least, for gin and tonic.

[3]Fish consumtion data are from 2007, retrieved from FAOSTAT. Coastline length and land area by country were obtained from The CIA Factbook.

relative to that of females and that induced men to specialize in 'market production' while women specialized in 'home production.' In turn, this specialization gave rise to beliefs and attitudes that have persisted until the present and manifest themselves in, for example, lower female labor force participation among first and second generation immigrants to the US from countries that historically used the plough.[4] Ross (2008) describes the case of oil producing countries and suggests that it is the fact that the oil producing technology crowds out women from the labor market what reinforces the prevalence of patriarchal norms and values. These two papers underscore the role of technology in shaping individual preferences, which are then shown to exhibit a significant correlation with, among other things, female labor market participation rate.[5]

Our paper focuses on the determination of labor market participation rate in a model in which preferences are shaped by a culture whose evolution is driven by technology.[6] We build a simple OLG model that captures the evolving nature of culture. Every (two period lived) generation will be identical except for the culture they inherit from the previous generation. For a given generation, the current state of culture and the economic decisions of the individuals given prevailing wages will determine their labor market participation rate in the second period of their lives. In the steady state, the labor market participation rate will be determined by the marginal product of labor (our technological parameter).

For each generation, as the labor market participation rate depends on the quit rate, the received culture will affect wages offered by firms (see e.g., Oi, 1962). Hence, we will have reciprocal causation between labor market participation rate and expected present value of wages. Note that we refrain from talking about wages and use instead the more cumbersome 'expected present value of wages' expression. This is because, although perfect competition among firms guarantees that the expected present value of labor costs must equal the expected present value of the marginal product of the workers, it has no direct implication on the particular value of the wages paid at every point in time. Therefore, some additional assumption regarding how each period wage is fixed is needed to close the model. Contract theory provides the natural option: we impose that firms offer whichever wage schedule is preferred by individuals.[7]

---

[4]Boserup (1970) is credited by Alesina et al. (2010) as the origin of the idea that it is the historic mode of agricultural production what shaped, at least in part, gender role attitudes.

[5]A related literature is that of papers dealing with the impact of external events on preferences. Examples are, Nunn and Wantchekon (2011), Alesina and Fuchs-Schündeln (2007) or Guiso et al. (2008). Nunn and Wantchekon (2011) trace back mistrust to the likelihood of ancestors being subject to slave trade. Alesina and Fuchs-Schündeln (2007) document the impact of Comunism on East Germans preferences. Guiso et al. (2007) relate differences in social capital between the North and South of Italy to the free city states experience in the North of Italy at the turn of the first millennium.

[6]We understand that culture need not be neutral with respect to technological change, and that, to some extent, changes in economic fundamentals may be driven by culture itself. However, we will concentrate in the economics to culture direction of causality.

[7]Some papers have indeed pursued this line of thought (e.g., Lazear and Rosen, 1990), but they have always

Our model, thus, integrates both contract theory and the role of culture in an OLG model to shed some light on the labor market participation decision. In this sense, our paper is related to Hauk and Saez-Marti (2002), Escriche et al. (2004) and Escriche (2007). In these papers, authors embed cultural transmission of values à la Bisin and Verdier (2001) and asymmetric information in an OLG framework. The main difference between our approach and theirs is that they all take as given the existence of several subcultures, each characterized by specific preferences, and focus on determining their equilibrium distribution, i.e., the proportion of individuals having each of those preferences. These papers, thus, feature an endogenous distribution of preferences but specific preferences themselves are unaffected by economic factors. In contrast, our agents are ex ante identical and it is only changes in the economic fundamentals what would give rise to the appearance of these differentiated groups.

In some sense, hence, it could be argued that our model rationalizes heterogeneous preferences: to the extent that individual preferences can be the product of other economic fundamentals, heterogeneity in preferences would simply reflect variation in the relevant fundamentals. For example, one could account for the male-female labor market participation gap as the result of past differences in productivity, as described in Alesina et al. (2010). In fact, heterogeneous preferences can arise from any factor that impacted the relevant technological parameter. For example, we would expect individuals living in areas in which property rights enforcement is limited to work shorter hours. This, in turn, might lead to more negative attitudes towards work. In this vein, the difference in hours worked across countries could be understood as taxation originating the preference for leisure that some (e.g. Blanchard, 2004) suggest as an explanation.

The paper is organized as follows. Section 2 presents the model for consumers and firms. Section 3 describes the equilibrium conditions. Section 4 analyzes some implications of the model and Section 5 concludes.

## 2 The Model

### 2.1 Consumers

Each generation is formed by a continuum of individuals (indexed by $i$) of measure 1 who live for two periods, say $t$ and $t+1$, and are endowed with one indivisible unit of labor time per period. Individuals derive utility from current and future consumption and from future leisure according to a time separable utility function

$$U_i\left(c_t, c_{t+1}, L_{t+1}\right) = u\left(c_t\right) + \beta u\left(c_{t+1}\right) + \beta\lambda_i\left(1 - L_{t+1}\right)$$

confined themselves to the static (or one generation) case.

where $c_t, c_{t+1} \in \mathbb{R}_+$ denote time $t$ and time $t+1$ consumption levels, $L_{t+1} \in \{0,1\}$ indicates labor time, $\beta$ is a subjective discount factor and $\lambda_i$ is the marginal valuation of time, which is unverifiable in the asymmetric information sense. This implies not only that markets are incomplete, but also that they cannot be completed. Since individuals do not value current leisure, they inelastically supply their endowment time to the labor market when young. At times it will be convenient to indicate whether we are referring to the generation an individual belongs to; in those cases, the superscripts $t$ and $t+1$ will be used, but, otherwise, they will be dropped to enhance readability. We assume further that $u(\cdot)$ is differentiable, strictly increasing and strictly concave, that all derivatives have constant sign over the entire domain and that the usual Inada conditions apply. We also assume that the absolute risk aversion, as measured by the Arrow-Pratt coefficient, is non increasing in income for $u(\cdot)$.

Individuals do value leisure in the second period of their lives, but only learn about their marginal valuation of leisure at the onset of that second period. The idea behind this specification is that, as the evidence cited by Bowles (1998) indicates, preferences are commonly formed as a result of the previous consumption experience. In our case, for preferences for leisure to develop, the individual must first experience work, which she does when young. This working experience allows her to develop a taste for leisure (or a distaste for work), upon which she will subsequently choose whether to work or not when old. Before working, individuals are only aware of the common (to all individuals) distribution function of leisure valuation. In particular the known distribution of $\lambda_i$ is as follows, let $q_{t+1}$ be the probability with which an individual born at time $t$ receives $\lambda_i = 0$ (we will refer to $q_{t+1}$ as the unconditional probability of working), and $F(\lambda)$ the conditional cdf of $\lambda$ for $\lambda_i > 0$. We assume that $F(\lambda)$ is twice continuously differentiable and its density, $f(\lambda)$, is strictly positive and non-increasing. The unconditional probability of working can be understood as reflecting the prevailing attitudes towards work and leisure, and, in this sense, as reflecting the impact of culture on labor market participation. From the point of view of each generation, this unconditional probability of working is given because the culture in which they are bred is heavily influenced by that of the previous generation.

Hence, consumers born at time $t$ only learn about their particular marginal valuation of leisure before time $t+1$ consumption and leisure or work takes place. This specification of uncertainty implies that all individuals are *ex-ante* (as of time $t$) homogeneous and *ex-post* (as of time $t+1$) heterogeneous and is akin to the one used in Lazear and Rosen (1990), whose workers only receive information about their reservation wage just before deciding whether to work or not. At the time of hiring, therefore, an individual has no more information about her characteristics than her potential employers and this makes the problem one of genuine uncertainty rather than one of asymmetric information. However, as there is no "law of large numbers" applicable in the context of a continuum of agents (see Feldman and Gilles, 1985), this individual uncertainty translates into aggregate uncertainty which we would like to avoid. We, thus, take the approach in Alós Ferrer (2002) whereby individual uncertainty is specified so that

7

it disappears in the course of aggregation. This specification removes aggregate uncertainty at the cost of losing independence between individuals, but this is immaterial in our model, because it only implies that the index $i$ is correlated with the realization $\lambda_i$. In so far as the index itself is not observable, this is irrelevant.

Consider an individual born at time $t$. On the first period of her life, the individual is assumed to work, and will receive a wage of $w_t$ that will be devoted to current consumption, $c_t$, and to savings, $s_t$, which will earn interest $R_{t+1}$. The decision on whether to work in the next period hinges upon the actual realization of the $\lambda_i$ parameter she receives. At time $t+1$ she will choose between staying employed, which allows consumption $\bar{c}_{t+1} = s_t R_{t+1} + w_{t+1}$, or quitting her job, in which case her consumption will be of only $\underline{c}_{t+1} = s_t R_{t+1}$ but she will receive $\beta \lambda_i$ in additional utility from leisure. It is clear that she will remain employed if and only if $\lambda_i < \widehat{\lambda}_{t+1} = u\left(\bar{c}_{t+1}\right) - u\left(\underline{c}_{t+1}\right)$, a critical value that is decreasing in savings and increasing in the wage received when old. Therefore, the probability with which an individual born at time $t$ will work at time $t+1$, will be, thus,

$$(1) \qquad\qquad p_{t+1} = q_{t+1} + \left(1 - q_{t+1}\right) F\left(\widehat{\lambda}_{t+1}\right).$$

It is important to note that our previous assumptions on the distribution of $\lambda_i$ have a number of implications. First, the probability of working will be always strictly below one. Second, the Inada conditions together with $0 < p_{t+1} < 1$ imply non-negative savings. And third, $p_{t+1}$ coincides with the proportion of old individuals that will work at time $t+1$. Culture, thus, determines a 'base labor market participation rate,' $q_{t+1}$, whose influence is combined with that of economic factors to yield the actual labor market participation rate, $p_{t+1}$.[8]

The problem for the consumer is, thus, to choose $s_t$ to maximize her expected utility of current and future consumption and of (contingent) future leisure, subject to her budget constraints, i.e., $c_t + s_t = w_t$ and $c_{t+1} = s_t R_{t+1} + w_{t+1} L_{t+1}$, and taking into account that the individual chooses $L_{t+1} = 1$ if and only if $\lambda_i < \widehat{\lambda}_{t+1}$. The objective function of the individual can be written as

$$E_{f(\lambda_i)}\left[U_i\left(c_t, c_{t+1}, L_{t+1}\right)\right]$$

$$= q_t\left[u\left(c_t\right) + \beta u\left(\bar{c}_{t+1}\right)\right] + \left(1 - q_t\right)\int_0^\infty \left\{u\left(c_t\right) + \beta\left[u\left(c_{t+1}\right) + \lambda_i\left(1 - L_{t+1}\right)\right]\right\} f\left(\lambda_i\right) d\lambda_i$$

$$(2) \quad = u\left(c_t\right) + E_{p_{t+1}}\left[\beta u\left(c_{t+1}\right)\right] + \beta\left(1 - q_t\right)\int_{\widehat{\lambda}_{t+1}}^\infty \lambda f\left(\lambda_i\right) d\lambda_i$$

$$(3) \quad = u\left(c_t\right) + E_{p_{t+1}}\left[\beta u\left(c_{t+1}\right)\right] + \left(1 - p_{t+1}\right)\beta E_{f(\lambda_i)}\left[\lambda_i \,\middle|\, \lambda_i > \widehat{\lambda}_{t+1}\right],$$

where $E\left[\cdot\right]$ and $E\left[\cdot\middle|\cdot\right]$ are the expectation and the conditional expectation operators respectively and the subindex indicates the probability under which the expectation is taken. The FOC can

---

[8] The time $t+1$ labor market participation rate is actually $\frac{1+p_{t+1}}{2}$.

be shown to be the usual

$$-u'(c_t) + E_{p_{t+1}} \left[ u'(c_{t+1}) \beta R_{t+1} \right] = 0. \tag{4}$$

To understand the disappearance of terms connected to leisure, observe that as the probability of working increases the individual gives up utility of leisure in exchange for utility from consumption. At the margin, the added probability of working, reduces the utility from leisure in exactly its critical value, i.e., $\widehat{\lambda}_{t+1}$; also at the margin, the gain in utility from consumption is $u(\bar{c}_{t+1}) - u(\underline{c}_{t+1})$. This two values are identical, and, hence, it is as if savings had no effect on the utility of leisure.[9]

Note that the Inada conditions imply that the LHS of (4) is (large) positive for values of $s_t$ close to 0 and (large) negative for values close to $w_t$, which together with its continuity guarantees existence of optimal savings. However, the effect on the LHS of (4) of an increase in savings is unclear because, in addition to the usual negative effect on the expected marginal utility of consumption, there is also a positive indirect effect that operates through the increase in the probability of working, as can be seen in

$$\frac{\partial^2 E_{f(\lambda_i)} \left[ U_i(c_t, c_{t+1}, L_{t+1}) \right]}{(\partial s_t)^2} = u''(c_t) + E_{p_{t+1}} \left[ \beta u''(c_{t+1}) R_{t+1}^2 \right]$$
$$+ \frac{\partial p_{t+1}}{\partial \widehat{\lambda}_{t+1}} \left( \frac{\partial \widehat{\lambda}_{t+1}}{\partial s_t} \right)^2.$$

To guarantee that a unique solution to this problem exists, we provide a natural extension of the traditional definition of 'normal good' to our uncertain environment: we characterize a good as 'generalized normal' if an increase in current wealth leads to an increase in its expected consumption.[10]

**Proposition 1** If time $t + 1$ leisure is a generalized normal good as defined, optimal savings are uniquely determined. In addition, they are non decreasing in $w_t$ and non increasing in $w_{t+1}$.

**Proof** The expected time $t + 1$ leisure is $E[1 - L_{t+1}] = 1 - p_{t+1}$, and the above definition of a generalized of normal good implies that

$$\frac{\partial E[1 - L_{t+1}]}{\partial w_t} = -\frac{\partial p_{t+1}}{\partial s_t} \frac{\partial s_t}{\partial w_t} > 0.$$

As the probability of working has been shown to decrease with savings, it must be the case that $\frac{\partial s_t}{\partial w_t} > 0$. The only way an individual can increase the expected amount of leisure

---

[9] From an analytical standpoint, this is more easily seen taking the derivative of (2) with respect to $s_t$, rather than that of (3).

[10] The following proposition can be viewed as an extension of the classical result in Diamond (1965).

she will experience is through the reduction of the probability of working, and this is accomplished by increasing the savings level. Note that this assumption does not imply necessarily that time $t+1$ consumption is also a normal good because, although the increase in savings will lead to an increase in $c_{t+1}$ irrespective of whether the individual works or not at $t+1$, it will also shift the probability towards the low consumption state and, hence, the net effect on the expected future consumption is unclear.

Implicit differentiation of the FOC yields

$$\frac{ds_t}{dw_t} = \frac{u''(c_t)}{\frac{\partial^2 E_{f(\lambda_i)}[U_i(c_t, c_{t+1}, L_{t+1})]}{(\partial s_t)^2}},$$

which has been shown to be positive. As the numerator is negative, this implies a negative denominator. This ensures that a unique strictly positive solution to (4) exists, and optimal savings $\widehat{s}_t^j = S(w_t, w_{t+1}, r_{t+1}, q_{t+1})$ are well defined. The amount saved depends, thus, on the wages paid at $t$ and $t+1$ and on the prevailing interest rate between these two periods. We have already shown $\frac{dS(w_t, w_{t+1}, R_{t+1})}{dw_t} > 0$, and to establish the sign of $\frac{dS(w_t, w_{t+1}, R_{t+1})}{dw_{t+1}}$, we calculate

$$\frac{\partial^2 E_{f(\lambda_i)}[U_i(c_t, c_{t+1}, L_{t+1})]}{\partial s_t \partial w_{t+1}} = E_{p_{t+1}}\left[\beta u''(c_{t+1}) R_{t+1}\right] + \frac{\partial p_{t+1}}{\partial w_{t+1}}\frac{\partial \widehat{\lambda}_{t+1}}{\partial s_t} < 0$$

i.e., consumers react to a higher future wage by reducing savings and, thus, increasing the probability of working.

## 2.2  Firms

We assume that there is perfect competition among a continuum firms, who share a common technology displaying constant returns to scale. As firms are assumed to be able to hire as much capital as they want at a constant cost $r$, they choose a fixed capital per worker ratio, which, in turn, determines that output per worker is constant. Of course, if output per worker, capital per worker and the cost of capital are all constant, the marginal product of labor will also be constant, and will be denoted $\overline{w}$. In addition, financial arbitrage implies that the rate of return on savings, $R_{t+1}$, must coincide with the cost of capital $r$. We also impose that all workers are equally productive, irrespective of their age or type, and that firms face a one-time cost when hiring a new worker. This cost is borne when hiring irrespective of the age of the worker, and, therefore, if an old worker wanted to work for any firm other than the one that hired her when young, the new firm would have to pay this cost. We assume this cost, $\kappa$, to be constant, although results are qualitatively unaffected if it is proportional to the wage.

Denote $P_{t+1}$ the proportion of old individuals willing to work at the prevailing wage; as

already noted, our previous assumptions on the distribution of $\lambda_i$ imply that $P_{t+1} = p_{t+1}$,[11] and, hence, $P_{t+1} = P(s_t, w_{t+1})$. We assume that firms do not face any uncertainty with respect to the measure of workers that will choose to stay on the job when old. To guarantee this, in view of the aforementioned inexistence of a law of large numbers for a continuum of agents, we impose that each firm hires a countably infinite number of workers. This assumption warrants that the ex-ante probability of working that each individual faces coincides with the ex-post proportion of old-age agents that choose to work not only for the overall economy but also at the firm level. In addition, for each young worker the firm employs at time $t$, it will also be able to employ $P_{t+1}$ old workers at time $t + 1$.[12]

Competition among firms ensures that firms make zero profits. For profits to be zero, the present value of the expected worklife labor cost of each worker type must coincide with the present value of her expected worklife output,

$$(5) \qquad w_t^t + \kappa + w_{t+1}^t \frac{P_{t+1}}{r} = \overline{w} \left( 1 + \frac{P_{t+1}}{r} \right).$$

In addition, $w_{t'}^t$ is bounded below

$$(6) \qquad w_{t'}^t \geq \overline{w} - \kappa, \quad t' = \{t, t + 1\},$$

i.e., no wage rate can fall below the $\overline{w} - \kappa$ threshold. Imagine $w_t^t < \overline{w} - \kappa$, this would make young workers 'too cheap' and a firm that only hired these workers (and fired them when old) would be able to make positive profits. The fact that old workers always receive at least $\overline{w} - \kappa$ is what warrants that they have no incentive to change firms when old and justifies our previous claim that firms face no uncertainty about the number of old workers it can employ. Of course, if the firm could credibly commit itself to time $t + 1$ wages for old workers, wages for young workers could fall short of $\overline{w} - \kappa$. Even in this case, however, the lower bound for the wage of old workers would be binding. To summarize, then, the contract space we are considering is the set of all wage pairs that satisfy (5) and (6) simultaneously. Clearly, the zero profits condition (5) only places a restriction on the present value of the wages paid to a worker, but not on the time distribution of this amount, i.e., not on the particular values of $w_t$ and $w_{t+1}$. The lower bounds for wages in (6) reduce further the set of admissible wage pairs. We argue that competition among firms leads them to offer whichever distribution is most desirable for workers, i.e., competition forces firms to choose $w_t, w_{t+1}$ so that individuals maximize their utility among those wage pairs for which profits are zero and are above the lower bound, i.e.,

$$(7) \qquad \{w_t, w_{t+1}\} \in \arg\max \{(3) \ s.t. (5), (6)\}$$

---

[11] This is tantamount to assuming that firms compute $P_{t+1}$ under the assumption that the continuum of individuals behave as would a representative (measure one) individual.

[12] For this statement to be accurate, we need that those individuals willing to work when old have no incentives to accept offers from other employers. We will see below that the conditions under which this is true are implied by the zero profits condition.

This condition is akin to the incentive compatibility constraints commonly found in the asymmetric information literature. However, it is worth stressing that in our setting it arises as a consequence of competition among firms, and not as a restriction on the problem of the firm.

There is an interesting specific case: when $w_t = \overline{w} - \kappa$ and $w_{t+1} = \overline{w}$, i.e., every worker is paid her current marginal product net of current hiring costs. We label this contract the 'age contract.' Under this contract, firms obtain the same profits from hiring any worker, irrespective of her age. Furthermore, wages do not depend on the quit rate. In all other contracts, however, wages and quit rates are jointly determined.

In all, competition among firms has three effects. First, firms make zero profits $(5)$. Second, all wages must be above a given threshold $(6)$. And third, within the limits imposed by the previous two conditions, the time distribution of wages is the one preferred by workers $(7)$.

# 3 Equilibrium

## 3.1 Equilibrium within a generation

Consider the generation born at time $t$ and recall that members of this cohort take $q_{t+1}$ as given. We study first the case where $q_{t+1} = q, \forall t$. For this generation, an equilibrium is a manifold that includes the savings decision, the probability of working and the wage rates, i.e., equations $(4)$, $(1)$, and $(7)$, must be satisfied simultaneously. Let $R_A(c_t)$ denote the Arrow-Pratt measure of absolute risk aversion, and $\lambda_m = u(\overline{w}(1+r) - \kappa) - u(r\overline{w})$. Proposition 2 establishes the existence of the equilibrium for such an economy.

**Proposition 2** If

$(2.i)$ $f(\lambda_m) \leq M,$[13] and

$(2.ii)$ time $t+1$ leisure is a generalized normal good

then there exists a unique, stable equilibrium for the economy.

Condition $(2.i)$ imposes a limit on the probability mass in the right tail (beyond some point $\lambda_m$) of the distribution of $\lambda_i$.

**Proof** We give here a sketch of the proof and defer the formal proof to Appendix A.1[14]. We treat the problem in two separate parts. First, by substituting (1) in (4), (5) and (7), we

---

[13]$M$ is a positive real number whose precise definition is given in Appendix A.1.

[14]In the Appendix, we ommit most technical details for the sake of brevity. Detailed proof is available from the authors.

construct the following value function, which is well defined given $s_t$ :

$$V\left(s_t\right) = \max_{\{w_t, w_{t+1}\}} u\left(c_t\right) + E_{p_{t+1}}\left[\beta u\left(c_{t+1}\right)\right] + \left(1 - p_{t+1}\right)\beta E_{f(\lambda)}\left[\lambda_i \left| \lambda_i > \widehat{\lambda}_{t+1}\right.\right]$$

$$s.t.\,(5)\,,(6)\,,\text{ and } w_t - s_t \geq 0$$

Under the assumptions, both the objective function and the restrictions are quasi-concave, and given that a solution to this problem exists and that the constraint qualification is met everywhere, the Kuhn-Tucker conditions are necessary and sufficient to characterize the unique solution to the problem, $\{w\left(s_t\right), p_{t+1}\left(s_t\right)\}$, with $w\left(s_t\right) = \{w_t, w_{t+1}\}$. Note that the inverse function theorem ensures that this solution is differentiable with respect to $s_t$. Second, we construct an auxiliary function $g\left(s_t, w\left(s_t\right)\right)$ that is simply the first order condition of the problem of the individual when wages are such that solve the problem of the firm, and show that it has at least one zero. Let $s^*\left(w_t, w_{t+1}\right)$ be such that $g\left(s^*, w\left(s^*\right)\right) = 0$. Then, $s^*$ defines an equilibrium for the economy. Note that, in principle, any feasible contract can be an equilibrium contract, i.e., equilibrium considerations do not limit further available contracts. This implies that the equilibrium contract can be, but need not be the age contract. Finally, we ensure that irrespective of the equilibrium contract, this equilibrium is unique.

### 3.1.1 Equilibrium wage schedules

The presence of the seniority payments in the labor market is pervasive. However, the age contract is the maximally back loaded wage scheme i.e., the one that implies the largest difference between young and adult wages (or the highest seniority payments). We are interested in determining under what conditions individuals prefer wage schedules involving a less steep wage schedule. Therefore, we find conditions on the primitives of our model that guarantee that the contract prevailing in equilibrium is not the age contract. If this is the case, as we have already noted, wage rates will be dependent on the quit rate.

Consider a wage schedule consisting in equal payments in both periods, i.e., $w_t^t = w_{t+1}^t$, and label it 'constant wage schedule.' We will prove that, under some conditions, it is strictly preferred to the age contract and that this preference is robust in the space of parameters for which proposition 2 holds.[15]  Obviously, if a constant wage schedule is preferred to the age contract, this constant wage schedule may or may not be the equilibrium contract but we are certain that the equilibrium contract will be distinct from the age contract. Let $\underline{\beta}, \underline{W}, \underline{r} > 0$ denote the minimum values $\beta, \overline{w}$, and $r$ can take, also let $\underline{s} = S\left(\overline{w} - \kappa, \overline{w}\right)$, and observe that $\underline{s}$ will be the minimum savings for given parameter values. Our initial parameter space is, then $\widehat{\Theta} \subset$

---

[15]This preference will, hence, be shown to be robust with respect to a particular $u\left(c\right)$. We take this to be given from the outset, since the range of the cdf, the density and the elasticities are dependent on its functional form, as well as on $\overline{w}$

$[\underline{\beta}, 1) \times [0, 1] \times (0, 1] \times [\overline{W}, \infty) \times [\underline{r}, \infty) \times \mathbb{R}_{++}$, with typical element $\widehat{\theta} = (\beta, q, \kappa, \overline{w}, r, f(\lambda_m))$, where $\lambda_m = u(\overline{w}(1 + r) - \kappa) - u(r\overline{w})$. Also note that, given the definition of the parameter space and the consumer problem, we can obtain $\inf_{\widehat{\theta} \in \widehat{\Theta}} \underline{s} = \underline{\underline{s}} > 0$, which can be taken as a parameter, as it is independent of the specific values of the other parameters. Hence, we take our enlarged parameter space as $\Theta = \widehat{\Theta} \times \mathbb{R}_{++}$, with typical element $\theta = (\beta, q, \kappa, \overline{w}, r, f(\lambda_m), R_A(r\underline{\underline{s}}))$.

**Proposition 3** Let $\mu(\cdot)$ be the Lebesgue measure on $\Theta$ and, given $\theta \in \Theta$, $w_t^*(\theta) = \{w_t^*, w_{t+1}^*\}$ be the unique equilibrium wage schedule for the generation born at time $t$. Then, there exists $I \subset \Theta$ such that $\mu(I) > 0$ and for all $\theta \in I$, $w_t^*(\theta) \neq \{\overline{w} - \kappa, \overline{w}\}$.

**Proof** We offer a formal proof in Appendix A.2 that proceeds along the following lines: We first construct a 'constant wage' contract, i.e., one in which $w_t^t = w_{t+1}^t$ that satisfies (5) and (6), and show that, for small values of $\kappa$, it is preferred to the age contract. As (7) has to be satisfied in equilibrium, this implies that the prevailing wage scheme will be different from the age contract.

In words, this implies that there is an open set of parameters for which the equilibrium of the economy will imply a wage schedule other than the age contract. An economic rationale for this is that, in presence of hiring costs, the constant wage contract provides some insurance against the eventuality of not working when old, because consumers receive higher income (compared to the age contract) when young (i.e., when they are sure to work), and, hence, are better able to smooth their consumption levels.

## 3.2 Steady state equilibrium

We have already mentioned that the unconditional probability of working reflects the impact of culture on labor market participation. In particular, we posit that individuals observe the quit behavior of the previous generation and derive from it some assessment of the 'acceptability' of quitting. On the other hand, it is already clear that the received culture is not the only relevant factor, because the economic environment is also crucial in the labor market participation rate determination. The inherited unconditional probability of working, together with the economic environment the individuals face, translates into their actual behavior regarding quits, i.e., the actual probability of working is determined according to (1). The fact that individuals do not simply work (or quit) with their inherited probability, but change it to account for economic factors is key, because it causes culture to evolve through time in response to these economic factors. We have, thus, that the current value of economic variables is heavily influenced by culture, but that, at the same time, it is economic fundamentals that shape culture in the long run. We introduce, thus, a process that governs the evolution of $q_t$ through time. Let

$$q_{t+1} = \pi(q_t, p_t) : [0, 1]^2 \to [0, 1]$$

14

be the law of motion (which we assume to be common knowledge) for the unconditional probability of working. Therefore, $q_{t+1}$ can be viewed as the result of cultural influences received by the individual. Knowledge of past values of the unconditional probability of working (the past acceptability of quitting) and of the actual proportion of old age workers helps determining how acceptable it is to quit. The higher these values, the least acceptable should be this behavior and, hence, the higher the base labor market participation rate, $q_{t+1}$. As a consequence, the law of motion should be increasing in both $q_t$ and $p_t$. In addition, and to guarantee that the steady state equilibrium is unique and stable, we will assume that $\pi(q_t, p_t)$ is continuously differentiable and that its partial derivatives are bounded above by $\psi$, a positive real number which will be defined in Appendix A.2. Finally, as $q_{t+1}$ is a probability, we impose that

$$(8) \qquad\qquad 1 > \pi(q_t, p_t) > 0, \forall (q_t, p_t) \in [0,1]^2$$

The assumed law of motion for $q_t$ embodies the idea that preferences are partly transmitted across generations. However, this transmission can take on several forms and steams from different sources, from which three are especially relevant. First, there is transmission of cultural values within the family, and it can be explicit (i.e., intended by the individual parents) or implicit (i.e., derived from the repeated exposition of the individual as a child to certain situations, from which endogenous modifications of tastes arise). Moreover, both types of intra family transmission interact, as the individual can observe as a child their parents actual behavior and contrast it to the values explicitly transmitted. Second, individuals are embedded in a broader social framework, e.g. the neighborhood, the school or the institutional and legal framework of the country they live in. This represents a constraint on their actions and an external (to the family) source of interaction, and sets another channel for preference formation. Finally, the individual as a child can observe directly the actual economic behavior of the previous generation as a group and this observation can also impact the preference generation mechanism.

Our modelization of the process is flexible enough to accommodate as particular cases some of the usual theoretical approaches on preference formation. For example, in Fernández (2007), individuals receive from the previous generation a prior with respect to the disutility of work and update that prior based on some private signal. In our model, $q_t$ would be the inherited prior, while the fact that parents did work or not could constitute the private signal. Cultural transmission by aspirations, à la de la Croix (1996), can be incorporated rather simply by choosing $\pi(q_{t-1}, p_{t-1}) = \alpha q_{t-1} + (1-\alpha)\beta p_{t-1}$ where $\alpha$ describes the degree of aspiration with respect to the past generation explicit norms and $\beta$ stands for the degree of aspiration with respect to the past generation actual economic behavior. It is worth noting that, in models of aspirations, the utility the younger generation obtains from consumption in the first period of their lives usually depends on the consumption experience of the previous generation when young. In our model, it is the distribution of the marginal valuation of leisure what is endogenous and determines the probability of working and the expected consumption of leisure. Of course, to the extent that

this probability is influenced by the behavior of the previous generation, aspirations are reflected in it. And, because the probability of working is involved, these aspirations also have an impact on the utility derived from the consumption of other goods.

Irrespective of what the particular preference transmission mechanism is, we want to ensure that a unique steady state exists in this economy, and that is the purpose of the next proposition.

**Proposition 4** If

$(4.i)$ $\frac{\partial R_A(c_t)}{\partial c_t} \leq 0, \forall c_t,$

$(4.ii)$ $f(\lambda_m) \leq M^*.$[16]

$(4.iii)$ time $t+1$ leisure is a generalized normal good

$(4.iv)$ $q_{t+1} = \pi(q_t, p_t) : [0,1]^2 \to [0,1]$ is continuously differentiable, with positive partial derivatives bounded above by $\psi$ and $(8)$,

then there exists a unique steady state equilibrium for the economy, with $q_t = q \in (0,1)$, $\forall t$. This equilibrium is globally stable, in the sense that, for any initial value of $(q_0, p_0)$ the economy will converge to that unique steady state. Moreover, the steady state value of the unconditional probability of working is increasing in $\overline{w}$.

**Proof** As before, we give a sketch of the proof and refer the reader to Appendix A.3. First, given $q_t \in [0,1]$, we can apply the proof of proposition 2 to obtain a unique solution to the problem of the generation born at time $t-1$. Note that part of this solution is $p_t = q_t + (1-q_t) F(\widehat{\lambda}_t)$, which allows us to write it as $p_t = p_t(q_t)$. Hence, for generation born at time $t$, it is clear that $q_t$ is the only relevant state variable, and we can write

$$q_{t+1} = \pi(q_t, p_t(q_t)) = \pi(q_t)$$

and note that our previous arguments imply that the solution is differentiable with respect to the state variable $q_t$. Thus, we can obtain an expression for the differential

$$\frac{dq_{t+1}}{dq_t} = \frac{d\pi(q_t)}{dq_t} = \frac{\partial \pi(q_t, p_t)}{\partial q_t} + \frac{\partial \pi(q_t, p_t)}{\partial p_t} \frac{dp_t}{dq_t}.$$

This, together with the definition of $\frac{dp_t}{dq_t}$, which can be obtained from the solution to the problem of the generation born at time $t$, and assumption $(4.iv)$ above, ensure that the total effect on $q_{t+1}$ is less than proportional to the change in $q_t$ so that for all $(q_{t-1}, p_{t-1}) \in [0,1]^2$ :

$$\sup_{q_t, \in [0,1]} \left| \frac{dq_{t+1}}{dq_t} \right| = \sup_{q_t, \in [0,1]} \left| \frac{d\pi(q_t)}{dq_t} \right| = \sup_{q_t, \in [0,1]} \left| \frac{\partial \pi(q_t, p_t)}{\partial q_t} + \frac{\partial \pi(q_t, p_t)}{\partial p_t} \frac{dp_t}{dq_t} \right| < 1.$$

Hence $\pi(q_t)$ is a contraction on a complete metric space, and there exists a unique fixed point $q_{t+1} = \pi(q_{t+1}) \in (0,1)$.

---

[16] $M^*$ is a positive real number whose specific definition is given in Appendix A.3.

# 4  Implications

How can our model help explain the observed differences in labor market participation rates across groups defined e.g., by gender or by nationality? To the extent that $p_t$ is the product not only of current economic conditions but also of the inherited culture, $q_t$, all factors determining $q_t$ are relevant in the explanation of the present labor market participation rate. Notably, past values of the marginal product of labor will be shaping current attitudes towards work provided that the economy is not in a steady state: Imagine a world in which, individuals belonged to one of two distinguishable groups, call them 'Bees' and 'Drones.' Suppose that, initially, both the marginal product of labor and the distribution of the marginal valuation of leisure was identical across the two groups. Assume that an exogenous technological innovation causes the marginal product of Bees to increase while that of Drones remains constant, $\overline{w}^B > \overline{w}^D$. This could be what happened with the introduction of the plough, as Alesina et al. (2010) suggest: the superior strength needed to handle the plough increased male productivity, but left female productivity unchanged. Assume also that the prevailing wage scheme is not the one associated with the age contract. As shown in Proposition 3 above, this will be the case at least for small values of $\kappa$. In such a world, the unconditional probability of working of each group would converge to different values, $q^B > q^D$. Hence, economic factors would give rise a "cultural" artifact that associates group membership with labor market attachment. In turn, this larger labor market attachment would imply a larger expected present value of wages for Bees than for Drones.

From this standpoint $\left(\overline{w}^B > \overline{w}^D,\ q^B > q^D\right)$ suppose now that some exogenous innovation caused the marginal product of Drones to increase and equal that of Bees, à la Galor and Weil (1996). Clearly, we would observe a transition phase during which both the expected present value of wages and the unconditional probability of working of Drones converge to the same values Bees experience. During this transition the expected present value of wages would be lower for Drones than for Bees even though they are equally productive, and we would observe a wage gap between them[17]. This wage gap would arise between groups of groups differing only in the culture they have inherited from the previous generation, and bears no relation whatsoever to education, productivity or any of the usual determinants of wage.

Frequently any difference between wages of distinguishable groups (e.g., male vs. female, or black or Mexican vs. white) that can not be ascribed to these 'usual determinants' is attributed to pure discrimination. Our model suggests that this need not be the case: differences in wages may arise as a result of pure cultural (as opposed to economic) factors, and, to the extent that they are not included in the analysis, discrimination estimates might be biased upwards. In this vein, Erosa, et al. (2005) show that the better part of the unexplained wage differential between

---

[17]Actually, we would observe a lifetime expected wage gap because it is the expected present value of wages what would be lower for drones than for bees. As the equilibrium wage schedule has not been determined, we make no claims about the existence of wage gaps between wages at each point in time.

men and women vanishes when one considers only women with no children. In our context, this would merely imply that women without children face cultural restrictions with respect to quitting similar to those men do.

Our model can also be brought to bear on the issue of cross country differences in labor market hours worked. There is broad consensus in this literature about the crucial role taxes play in explaining these differences (see Prescott, 2004), but explanations stressing differences in preferences over consumption/leisure choices are not uncommon (e.g. Blanchard, 2004). We suggest what could be viewed as a consensual alternative between the two previous explanations. Imagine two identical countries, both in steady state equilibrium. Being identical, all variables, e.g., wages or the unconditional probability of working, would take the exact same values. Now suppose one country introduces some form of taxation and spends the proceeds in a manner that those taxed also reap the benefits of spending. There is a wedge, however, between taxes paid and services received, perhaps due to collection costs. This wedge is what will be crucial. As a consequence of the wedge, individuals will work less, and this will generate a departure in the culture of the high tax country from that of the low tax country: on average, leisure will be more highly valued in the high tax country. Thus, we would observe that individuals work less in countries with higher taxes (or more precisely, with higher inefficiency in their tax collection mechanisms). This lower number of individuals working would be a consequence of both, taxes and the pro leisure culture that taxes would induce.

In addition, the model can also be use for policy evaluation: For example, Alesina et al. (2007) raised the issue of gender based taxation, essentially on efficient taxation grounds. They contend that, given the different labor supply elasticities of male and female workers, marginal tax rates should be different for this two groups in order to satisfy the Ramsey criterion. They argue that as a by product of achieving efficiency, gender based taxation would help close the gender income gap, because it would induce women (men) to work longer (shorter) hours. Obviously, the gender wage gap would also reduce on an after tax base, but the gross (pre-tax) gender wage gap would not be affected.

Our model suggests that such gender based taxation will have implications on quit rates which, in turn, will impact not only net, but gross wages. These 'induced' changes will reinforce the effect of the differential tax rates, provided the taxed and subsidized populations are chosen carefully. In particular, a small tax on old working males would, in addition to reducing their net wages, induce an increase in their quit rate, which, in turn, would further reduce their pre-tax expected present value of wages. If the proceeds of the tax were distributed among old working women, their net income would increase; this would cause their quit rate to decrease, and this lower quit rate would lead to an increase in their pre-tax expected present value of wages. Note that, although in Alesina et al. (2007) wages themselves are unaffected, in our model wages change in response to changes in quit rates. What does change in Alesina et al. (2007) are the

labor supply elasticities of male and female workers and they suggest that in the event of those elasticities converging, there would disappear the rationale for a differential tax treatment of men and women. In a similar vein, our results suggest that gender based taxation can be used to speed the convergence of quit rates. A formal treatment of this problem can be found in Appendix A.4.

# 5    Summary and conclusions

In recent years, we have witnessed a rapid and steady growth in papers dealing with the relationship between culture and economics, and in particular, with the influence of culture on preferences. Several strands of this literature can be identified, each of dealing with a particular aspect of the problem. There are papers that document (but do not attempt to model) changes in preferences stemming from changes in either the cultural environment or the economic fundamentals. There are also papers that emphasize the importance of the cultural background in explaining differences in behavior, but ignore what caused those diverse preferences. Finally, and somewhat independently from the previous literature, there is a number of papers dealing with the mechanism of intergenerational transmission of preferences. We build a model that integrates all these (to some extent partial) analysis. Individual preferences are shaped by culture, which, in turn, evolves through time in response to changes in the economic fundamentals.

In particular, and in keeping with a large part of this literature, we analyze the labor market participation rate, and we do so through the quit to non participation decision. Individuals may exit the labor force in the second period of their lives (thus determining the labor market participation rate) and, in taking this decision, they weight their marginal valuation of leisure, whose distribution reflects the inherited culture, against the economic incentives of working. The comparison is not trivial: as numerous empirical studies have shown, quit rates and wage rates depend on each other, and the model reflects it. In addition, it must be noted that even if the present value of the product of labor was fixed, its time distribution (the wages paid to young and old members of the same generation) would be not, and we impose that competition among firms leads them to offer whatever wage schedule workers prefer. Hence, at each time, the labor participation rate, and the wage schedule need to be determined simultaneously. The following generation observes the behavior with respect to labor market participation of its predecessors and updates its beliefs about the acceptability of quitting, thus modifying the distribution of the marginal valuation of leisure, i.e., the culture. This updating is done in a manner compatible with the mechanisms of intergenerational transmission of preferences most frequent in the literature. Culture, therefore, evolves through the generations and ends up reflecting the current state of economic fundamentals. In the transition, however, both past and current fundamentals are relevant in the determination of equilibrium values. This mechanism helps explain why it is

possible to observe differences in the present expected value of wages of workers even when they are equally productive, or why individuals in countries with higher taxes seem to develop a preference for leisure.

# References

[1] Abel, A., "Asset Prices under Habit Formation and Catching up with the Joneses," *The American Economic Review* 80 Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association (May 1990), 38-42

[2] Alesina, A., and N. Fuchs-Schündeln "Goodbye Lenin (or Not?): The Effect of Communism on People's Preferences," *The American Economic Review* 97 (September 2007) 1507–1528.

[3] Alesina, A., P. Giuliano, and N. Nunn, "*The Origins of Gender Roles: Women and the Plough*," mimeo, Harvard University, 2010.

[4] Alesina, A., A. Ichino and L. Karabarbounis, "*Gender Based Taxation and the Division of Family Chores*," NBER Working Paper No. 13638, November 2007.

[5] Alós-Ferrer, C., "*Individual Randomness in Economic Models with a Continuum of Agents,*" mimeo, Universität Konstanz, 2002

[6] Becker, G. S. and K. M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 96 (August 1988), 675-700.

[7] Bisin, A. and T. Verdier, "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory* 97 (April 2001) 298–319.

[8] Blanchard, O., "The Economic Future of Europe," *Journal of Economic Perspectives* 18 (Fall 2004), 3-26.

[9] Boserup, E., *Woman's Role in Economic Development*. (London: George Allen and Unwin, 1970).

[10] Bowles, S., "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions," *Journal of Economic Literature* 36 (March 1998), 75-111

[11] Constantinides, G. M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy*, 98 (Jun., 1990), 519-543.

[12] de la Croix, D., "The dynamics of bequeathed tastes," *Economics Letters* 51 (October 1996), 89-96.

[13] Diamond, P., "National Debt in a Neoclassical Growth Model," *American Economic Review* 55 (December 1965), 1126-1150.

[14] Duesenberry, J. S., *Income, Saving, and the Theory of Consumer Behavior*. (Cambridge, Mass.: Harvard Univ. Press, 1949).

[15] Erosa, A., L. Fuster and D. Restuccia, "A Quantitative Theory of the Gender Gap in Wages," University of Toronto Working Papers tecipa-199, September 2005.

[16] Escriche, L., "Persistence of occupational segregation: The role of intergenerational transmission of preferences," *The Economic Journal* 117 (April 2007), 837–857.

[17] Escriche, L. Olcina, G. and R. Sánchez, "Gender discrimination and intergenerational transmission of preferences," *Oxford Economic Papers* 56 (July 2004), 485-511.

[18] Feldman, M. and C. Gilles, "An Expository Note on Individual Risk without Aggregate Uncertainty," *Journal of Economic Theory* 35 (February 1985), 26-32.

[19] Fernández, R., "Culture as Learning: The Evolution of Female Labor Force Participation over a Century," NBER Working Paper No. 13373, September 2007.

[20] Fernández, R., "Does culture matter?," CEPR, Discussion Paper No. 7965, (August 2010).

[21] Galor O. and D.N. Weil, "The Gender Gap, Fertility, and Growth," *The American Economic Review* 86 (June 1996), 374-87.

[22] Guiso, L., P. Sapienza and L. Zingales, "Long Term Persistence," NBER working paper 14278, August 2008.

[23] Hauk, E. and M. Saez-Marti, "On the Cultural Transmission of Corruption," *Journal of Economic Theory* 107 (December 2002), 311–335.

[24] Lazear, E.P. and S. Rosen, "Male-Female Wage Differentials in Job Ladders," *Journal of Labor Economics* 8 (January 1990), S106-S123.

[25] Marshall, A., *Principles of Economics: An Introductory Volume.* 8th ed. (London: Macmillan, 1920).

[26] Nunn, N. and L. Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," *The American Economic Review* (2011, forthcoming).

[27] Oi, W.Y., "Labor as a Quasi-Fixed Factor," *The Journal of Political Economy* 70 (December 1962), 538-555.

[28] Prescott, E.C., "Why Do Americans Work So Much More Than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review* 28 (July 2004), 2–13.

[29] Ross, M.L., "Oil, Islam, and Women," *American Political Science Review* 102 (February 2008), 107-23.

[30] Stigler, G.J. and G.S. Becker, "De gustibus non est disputandum," *The American Economic Review* 67 (March 1977), 76-90.

[31] Stokey, N. and R. Lucas, *Recursive methods in Economic dynamics*, With Edward C. Prescott (Cambridge, Mass. and London: Harvard University Press, 1989)

# A  Appendix

## A.1  Proof of proposition 2.  Existence and uniqueness of the equilibrium within a generation.

Let $q_t = q \in [0,1], \forall t$ be given. As we are considering each generation separately, we will only keep the time subscript to distinguish wages earned when young from those received when old. In particular, $c$ will denote young age consumption while old age consumption will be either $\overline{c}$ or $\underline{c}$ in the event of working or not working when old respectively, while $p$ will denote the probability of working when old. Let $w_t^{\max}$ be the maximum wage the firm can offer to young workers when $w_{t+1}$ is set equal to its minimum value. As optimal savings have been shown to depend positively on the current wage and negatively on the (expected) future wage, the maximum amount saved will be $\overline{s} = S(w_t^{\max}, \overline{w} - \kappa)$, whereas minimum savings are positive and given by $\underline{s} = S(\overline{w} - \kappa, \overline{w})$.

Recall that we defined $\underline{\beta}, \underline{W}, \underline{r} > 0$ as the minimum values $\beta, \overline{w}$, and $r$ can take. Our initial parameter space is, then $\widehat{\Theta} \subset [\underline{\beta}, 1) \times [0,1] \times (0,1] \times [\overline{W}, \infty) \times [\underline{r}, \infty) \times \mathbb{R}_{++}$, with typical element $\widehat{\theta} = (\beta, q, \kappa, \overline{w}, r, f(\lambda_m))$, where $\lambda_m = u(\overline{w}(1+r) - \kappa) - u(r\overline{w})$[18]. Let $M$ denote the normalized value of the upper bound of the density function at $\lambda_m$, whose precise value will be given later. Also note that, given the definition of the parameter space and the consumer problem, we can obtain $\inf_{\theta \in \Theta} \underline{s} = \underline{\underline{s}} > 0$, which can be taken as a parameter, as it can always be computed and is independent of the values of the other parameters.[19] Also, we can define maximum probability $\overline{p} = q + (1-q) F\left(u\left(\overline{w} + r\underline{\underline{s}}\right) - u\left(r\underline{\underline{s}}\right)\right)$ Hence, we take our enlarged parameter space as $\Theta = \widehat{\Theta} \times \mathbb{R}_{++}$, with typical element $\theta = \left(\beta, q, \kappa, \overline{w}, r, f(\lambda_m), R_A(r\underline{\underline{s}})\right)$. Finally, let $\Lambda = u\left(\overline{w} + r\underline{\underline{s}}\right) - u\left(r\underline{\underline{s}}\right)$ be the maximum value that $\widehat{\lambda}$ can attain (that this is indeed a maximum can be seen from the definition of $\underline{\underline{s}}$ given above and the fact that $\widehat{\lambda}$ is decreasing in $s$). Before stating the proof, let us highlight that condition $(ii)$ holds true, for example, for any rescaled exponential density, say $f(\lambda) = J_1 e^{-J_2\lambda}$, $J_1, J_2 > 0$, where both parameters are chosen appropriately.

The strategy of the proof is based on solving first a problem that determines the firm's optimal wage offer under equilibrium conditions, for every given possible individual choice of savings. Then, we proceed to find that choice of savings that is a best response to itself in the sense that it is the optimal choice for the individual when presented with wages given precisely by the optimal wage offer of the firm for that savings level. That savings level determines an

---

[18] Since $\widehat{\lambda} = u(\overline{c}) - u(\underline{c})$ is increasing in $w_{t+1}$ and decreasing in $s_t$, the minimum value it can achieve in the set $\Omega(s)$ defined below is $\lambda_M$. Moreover, since $\frac{\partial f(\lambda)}{\partial \lambda} < 0$, by assumption, any upper bound for $f(\lambda_M)$ is also an upper bound for $f(\lambda)$ in the set $\Omega(s)$

[19] This can be seen by noting that for the case $\beta = \underline{\beta}, \overline{w} = \overline{W}, r = \underline{r}, q = 0, f(\lambda_M) = 0$, $\underline{s}$ is still defined by the FOC of the consumer and is strictly positive.

equilibrium.

The proof is organized in several steps, that are numbered to enhance readability:

1. We determine the optimal solution for the firm in equilibrium

   (a) Since the equilibrium level savings should allow firms to make non-negative prof-
   its, pay wages above the given thresholds and individuals to obtain strictly positive
   consumption when young, we compute the maximum savings $\widehat{s}_0$ for which this is
   possible.

   (b) For each possible $0 < s < \widehat{s}_0$, we define an ancillary problem, $V(s)$, that characterizes
   the firms optimizing behaviour in any equilibrium. This problem consists in taking as
   given the individual savings and choosing the optimal wage schedule that maximizes
   the individual utility subject to obtaining non-negative profits and satisfying the lower
   bounds for both wages.

2. We define $g(s)$ as the FOC of the individual when wages are given by $w(s)$. Observe that
   a zero of $g(s)$ completely characterizes optimizing behaviour by the individual when firms
   optimize in equilibrium. Therefore, we define an equilibrium as that level of savings, $s^*$,
   such that $g(s^*) = 0$,. We prove that $s^*$ exists and that it must satisfy $\underline{s} \le s^* \le \overline{s}$.

3. To guarantee uniqueness, we ensure that $g(s)$ is decreasing around $s^*$ for any $s^* \in (\underline{s}, \overline{s})$.

   (a) We write the condition $\frac{dg(s^*)}{ds} < 0$ as an inequality of the form $A > f\left(\widehat{\lambda}\right) B$.

   (b) We specify an upper bound on $f(\lambda_m)$ involving only primitives that guarantees that
   $g(s)$ is decreasing around $s^*$. Since the assumptions guarantee that the primitives
   satisfy such a uniform upper bound, this completes the proof.

### A.1.1 Firms optimal behaviour in equilibrium

**1.a** It is clear that if firms take savings as exogenous, they cannot achieve non negative profits
for all savings levels, because the non negativity of consumption in the first period requires that
$w_t \ge s$. We need to determine that upper bound. Imagine firms offered $w_{t+1} = \overline{w} - \kappa$, the
lowest possible amount. This would enable them to pay $w_t^{\max}$, the largest possible amount in $t$
(because $s$ is given) that produces zero profits. Obviously, the largest savings compatible with
this would be $\widehat{s}_0 = w_t^{\max}$ and this value is determined by :

$$a(\widehat{s}_0) = \overline{w}\left[\kappa\left(\frac{q + (1-q) F\left(u\left(\overline{w} - \kappa + \widehat{s}_0 r\right) - u\left(\widehat{s}_0 r\right)\right)}{r}\right)\right] - (\widehat{s}_0 + \kappa) = 0.$$

Note that the existence of $\widehat{s}_0$ is guaranteed by the fact that the above function is monotonic
in $s$. Moreover, note that such $\widehat{s}_0$ could never be optimal for the individual, since her first

period marginal utility of consumption would be infinite. Hence, since $\bar{s} = S\left(w_t^{\max}, \overline{w} - \kappa\right)$ is the optimal response of the individual when facing the wage schedule $\left(w_t^{\max}, \overline{w} - \kappa\right)$, it follows that $\bar{s} < \widehat{s}_0$.

**1.b** Given $s \in (0, \widehat{s}_0)$, let

$$V(s) = \max_{\{w_t, w_{t+1}\} \in A} E_{f(\lambda_i)}\left[U_i\left(c_t, c_{t+1}, L_{t+1}\right)\right]$$

s.t.

(9) $\qquad \left(\overline{w}\left(1 + \frac{p}{r}\right) - w_t - \kappa - w_{t+1}\frac{p}{r} \geq 0\right) \quad (\nu_1)$

(10) $\qquad \left(w_t - \overline{w} + \kappa \geq 0\right) \quad (\nu_2)$

(11) $\qquad \left(w_{t+1} - \overline{w} + \kappa \geq 0\right) \quad (\nu_3)$

(12) $\qquad \left(w_t - s \geq 0\right) \quad (\nu_4)$

be the problem whose solution gives the optimal behaviour of the firm in equilibrium, where $A = (0, +\infty) \times (0, \overline{w} + 1)$ is open and $\nu_i$ denote the corresponding Kuhn-Tucker multipliers. It can be shown that $V(s)$ is well defined, has a unique solution completely characterized by the Kuhn-Tucker conditions provided

(13) $$\beta\frac{F(\lambda_m)}{\Lambda}\left(\overline{w} - \kappa\right) R_A\left(\overline{w}\left(1 + r\right)\right) = M_1 > f(\lambda_m).$$

## A.1.2 Equilibrium savings: existence

Now, in order to ensure that the individual chooses $s$ through maximizing behavior, we construct the following function

$$g(s) = -u'\left(w_t(s) - s\right) + \beta r\left[pu'\left(w_{t+1}(s) + sr\right) + (1 - p)u'(sr)\right].$$

which is simply the FOC of the consumer evaluated at $s$ when wages are given by $w(s)$ Under the assumptions, $\lim_{s \to 0^+} g(s) = +\infty$, $\lim_{s \to \widehat{s}_0} g(s) = -\infty$, (since $\lim_{s \to \widehat{s}_0} w_t(s) = s$) and the continuity of $g(s)$ ensures that there exists $s^* \in (0, \widehat{s}_0)$ such that $g(s^*) = 0$. Note that such $s^*$ determines an equilibrium for this economy, as defined in the proposition. Moreover, $s^* \in [\underline{s}, \bar{s}]$. To see it, recall that $s^*$ being an equilibrium implies that it is an optimal response of the consumer to $w(s^*)$, that is $s^* = S\left(w_t(s^*), w_{t+1}(s^*)\right)$ and $w(s^*)$ being a solution to $V(s)$ implies that $\left(\overline{w} - \kappa\right) \leq w_t(s^*) \leq w_t^{\max}$, $\left(\overline{w} - \kappa\right) \leq w_{t+1}(s^*) \leq \overline{w}$. Then, since $S(w_t, w_{t+1})$ has been shown to be increasing in $w_t$. and decreasing in $w_{t+1}$, we have that $\underline{s} = S\left((\overline{w} - \kappa), \overline{w}\right) \leq s^* = S\left(w_t(s^*), w_{t+1}(s^*)\right) \leq \bar{s} = S\left(w_t^{\max}, (\overline{w} - \kappa)\right)$.

### A.1.3  Equilibrium savings: uniqueness

**3.a**  In order to ensure the uniqueness of this equilibrium, it is enough to show that for any $s^*$ in the open interval $(\underline{s}, \overline{s})$, $\frac{dg(s^*)}{ds} < 0$. Note that for any such $s^*$, $(10)$, $(11)$ and $(12)$ must not be binding (so that $\nu_2 = \nu_3 = \nu_4 = 0$ ) while $(9)$ must hold with equality. This can be seen as follows: as shown above, $(9)$ must hold with equality and $(12)$ is an strict inequality at any solution of the $V(s)$ problem. For $(10)$, note that if $w_t = (\overline{w} - \kappa)$, $(9)$ implies that $w_{t+1} = \overline{w}$, which is precluded by the fact that we are considering $s^* > \underline{s} = S((\overline{w} - \kappa), \overline{w})$. An analogous argument shows that $\overline{s}$ is the only potential steady state savings where $(11)$ holds with equality. Hence, $(11)$ must not be binding at any $s^* < \overline{s}$. It can be shown that $w(s)$ is a continuously differentiable function around any equilibrium savings, $s^*$.

Let $c^* = w_t - s^*$, $\overline{c}^* = w_{t+1} + s^*r$, $\underline{c}^* = s^*r$ and write

$$(14)\quad \frac{dg(s^*)}{ds} = -u''(c^*)\left[\frac{dw_t(s^*)}{ds} - 1\right] + \beta r(1-q)f\left(\widehat{\lambda}\right)\left[\frac{\partial\widehat{\lambda}}{\partial s} + u'(\overline{c}^*)\frac{dw_{t+1}(s^*)}{ds}\right]\frac{\partial\widehat{\lambda}}{\partial s}\frac{1}{r}$$
$$+\beta r\left[pu''(\overline{c}^*)\left(\frac{dw_{t+1}(s^*)}{ds} + r\right) + (1-p)u''(\underline{c}^*)r\right].$$

We want to show that $g(s)$ is decreasing around $s^*$, and we will prove the following stronger condition[20]

$$(15)\qquad\qquad \frac{dg(s^*)}{ds} - \beta r^2(1-p)u''(\underline{c}^*) < 0.$$

Changes in savings can be shown to affect optimal wages as follows

$$\left.\frac{dw_{t+1}}{ds}\right|_{s=s^*} = \frac{u''(c^*)b(1+rb)}{C_1 + b^2u''(c^*)} - \frac{rC_2 + u''(c^*)rb^2}{C_1 + b^2u''(c^*)}$$
$$- \frac{u''(c^*)\frac{b}{r}(\overline{w} - w_{t+1})(1-q)\left[u'(\overline{c}^*) - u'(\underline{c}^*)\right]}{C_1 + b^2u''(c^*)}f\left(\widehat{\lambda}\right)$$
$$= W_1 + W_2 + W_3f\left(\widehat{\lambda}\right),$$

where $b = \frac{1}{r}\left[(\overline{w} - w_{t+1})(1-q)f\left(\widehat{\lambda}\right)u'(\overline{c}^*) - p\right]$, $C_1$ and $C_2$ are given by

$$C_1 = \beta pu''(\overline{c}^*) + (1-q)\beta f\left(\widehat{\lambda}\right)(u'(\overline{c}^*))^2$$
$$+ (1-q)\frac{\nu_1}{r}\left\{(\overline{w} - w_{t+1})\left[f'\left(\widehat{\lambda}\right)u'(\overline{c}^*)^2 + f\left(\widehat{\lambda}\right)u''(\overline{c}^*)\right] - 2f\left(\widehat{\lambda}\right)u'(\overline{c}^*)\right\}$$
$$C_2 = \beta pu''(\overline{c}^*) + (1-q)\beta f\left(\widehat{\lambda}\right)\frac{\partial\widehat{\lambda}}{\partial s}\frac{1}{r}u'(\overline{c}^*)$$
$$+ (1-q)\frac{\nu_1}{r}\left\{(\overline{w} - w_{t+1})\left[f'\left(\widehat{\lambda}\right)\frac{\partial\widehat{\lambda}}{\partial s}\frac{1}{r}u'(\overline{c}^*) + f\left(\widehat{\lambda}\right)u''(\overline{c}^*)\right] - f\left(\widehat{\lambda}\right)\frac{\partial\widehat{\lambda}}{\partial s}\frac{1}{r}\right\},$$

---

[20]This condition will be useful in Appendix A.3

and the definitions of $W_i$ are obvious from the expression above. Note that $W_1, W_3 < 0$ and that both are bounded[21]. Let $\widehat{W}_1 = \inf_{\theta \in \Theta} W_1$, and $\widehat{W}_3 = \inf_{\theta \in \Theta} W_3$. Also

$$
\begin{aligned}
\left. \frac{dw_t}{ds} \right|_{s=s^*} &= -\frac{brC_2 - u''(c^*)\, b^2}{C_1 + b^2 u''(c^*)} + \frac{C_1\,(\overline{w} - w_{t+1})\,(1-q)\,\frac{\partial \widehat{\lambda}}{\partial s}\frac{1}{r}}{C_1 + b^2 u''(c^*)} f\left(\widehat{\lambda}\right) \\
&= A_1 + A_2 f\left(\widehat{\lambda}\right)
\end{aligned}
$$

where, again, the definitions of $A_i$ are obvious. Note that $A_2 < 0$. Using these values, we can now write $\frac{dg(s^*)}{ds}$ as

$$
\begin{aligned}
\text{(16)} \qquad \frac{dg\,(s^*)}{ds} &= u''(c^*)\,[1 - A_1] + \beta r^2 \left[ pu''(\overline{c}^*) + (1-p)\,u''(\underline{c}^*) \right] + \beta r\,[W_1 + W_2]\,pu''(\overline{c}^*) \\
&\quad + f\left(\widehat{\lambda}\right) \left\{ (1-q)\,\beta \left( \frac{\partial \widehat{\lambda}}{\partial s} \right)^2 + \beta r\,[W_1 + W_2]\,(1-q)\,\frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right. \\
&\quad \left. + W_3 \beta r \left[ pu''(\overline{c}^*) + (1-q)\,\frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right] - u''(c^*)\, A_2 \right\}.
\end{aligned}
$$

Expression (15) can be written, using (16), as

$$
\begin{aligned}
&- u''(c^*)\,[1 - A_1] - \beta r^2 pu''(\overline{c}^*) - \beta r\,[W_1 + W_2]\,pu''(\overline{c}^*) > \\
&\qquad f\left(\widehat{\lambda}\right) \left\{ (1-q)\,\beta r^2 \left( \frac{\partial \widehat{\lambda}}{\partial s} \right)^2 + \beta r\,[W_1 + W_2]\,(1-q)\,\frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right. \\
&\qquad \left. + \beta r W_3 \left[ pu''(\overline{c}^*) + (1-q)\,\frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right] - u''(c^*)\, A_2 \right\}.
\end{aligned}
$$

**3.b** It can be shown that the LHS of (**??**) is positive and

$$
-u''(c^*)\,[1 - A_1] - \beta r^2 pu''(\overline{c}^*) - \beta r\,[W_1 + W_2]\,pu''(\overline{c}^*) > \frac{-u''(c^*)\,[1 - \overline{p}]^2}{1 + \frac{1}{\beta r}\frac{R_A(r\underline{s})}{R_A(\overline{w}(1+r))}}.
$$

Therefore, (**??**) would follow if either the RHS is negative or if it is positive and

$$
\begin{aligned}
\text{(17)} \qquad \frac{-u''(c^*)\,[1 - \overline{p}]^2}{1 + \frac{1}{\beta r}\frac{R_A(r\underline{s})}{R_A(\overline{w}(1+r))}} \; &> \\
&f\left(\widehat{\lambda}\right) \beta r\,(1-q) \left\{ r \left( \frac{\partial \widehat{\lambda}}{\partial s} \right)^2 + [W_1 + W_2]\,\frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right. \\
&\quad \left. + W_3 \left[ \frac{pu''(\overline{c}^*)}{(1-q)} + \frac{\partial \widehat{\lambda}}{\partial s} u'(\overline{c}^*) \right] - \frac{u''(c^*)\, A_2}{\beta r\,(1-q)} \right\},
\end{aligned}
$$

---

[21] It is easily seen that these bounds involve only parameters.

Now, dividing both sides of (17) by $u'(c^*)$, dividing and multiplying $\frac{\partial \widehat{\lambda}}{\partial s}$ by $\widehat{\lambda}$ and using the definition of absolute risk aversion, we can rewrite it as

$$\left[\frac{1}{1+\frac{1}{\beta r}\frac{R_A(r\underline{s})}{R_A(\overline{w}(1+r))}}\right]\frac{R_A(c^*)[1-\overline{p}]^2}{C_3} > f\left(\widehat{\lambda}\right),$$

where

$$
\begin{aligned}
C_3 &= (1-q)\beta r^2 \left\{ \frac{\widehat{\lambda}^2 R_A(\widetilde{w}+s^*r)^2}{u'(c^*)} - [W_1+W_2]\frac{u'(\overline{c}^*)}{u'(c^*)}\widehat{\lambda}R_A(\widetilde{w}+s^*r) \right. \\
&\left. -W_3\frac{u'(\overline{c}^*)}{u'(c^*)}\left[\frac{p}{(1-q)r}R_A(\overline{c}^*)+\widehat{\lambda}R_A(\widetilde{w}+s^*r)\right]+\frac{R_A(c^*)A_2}{\beta r^2(1-q)} \right\}
\end{aligned}
$$

Now, note that:

$i$) since $c^* < \overline{c}^* < \overline{w}(1+r)$ , we have that $\frac{u'(\overline{c}^*)}{u'(c^*)} < 1$ and $\frac{u'(\overline{w}(1+r))}{u'(c^*)} < 1$.

$ii$) It can be shown that $-\beta r[W_1+W_2] < -\beta r[W_1-r] < -\beta r\left[\widehat{W}_1-r\right]$ and, by the definitions given above, we have $-W_3 < -\widehat{W}_3$

$iii$) The term $R_A(c^*)A_2$ is negative.

$iv$) $\lambda_m$ is a lower bound for any equilibrium $\widehat{\lambda}$.

$v$) $\Lambda$ is the maximum value that $\widehat{\lambda}$ can attain.

$vi$) By the definition of $\underline{s}$ given above, for any $\widetilde{w} \in (0, w_{t+1})$, $s^*$ such that $g(s^*)=0$, we have $R_A(\widetilde{w}+rs^*) < R_A(rs^*) < R_A(r\underline{s}) < \infty$ .

Thus, using the appropriate (either the largest or smallest) values for all these variables in $C_3$, results in

$$\overline{C}_3 = R_A(r\underline{s})(1-q)\beta r^2 \left\{ \frac{\Lambda^2}{u'(\overline{w}(1+r))}R_A(r\underline{s}) - \left[\widehat{W}_1-r\right]\Lambda - \widehat{W}_3\left[\frac{1}{(1-q)r}+1\right]\right\}$$

being the largest possible value $C_3$ can take. Hence, a sufficient, but not necessary, condition for (17) to hold is

(18)
$$\left[\frac{1}{1+\left[\frac{1}{\beta r}\right]\frac{R_A(r\underline{s})}{R_A(\overline{w}(1+r))}}\right]\left[\frac{R_A(c^*)[1-\overline{p}]^2}{\overline{C}_3}\right] = M_2 > f(\lambda_m).$$

Note that $M_2$ does not depend on any endogenous object. Now we can define $M = \min[M_1, M_2]$. Since assumption $(ii)$ guarantees that (13) and (18) hold, this completes the proof.[22]

---

[22] Note that for the case $q=1$, we obtain $p_t=1$. In that case the problem is reduced to one without uncertainty, and thus our basic assumptions on the utility function ensure the uniquenness of the result.

## A.2 Proof of proposition 3: Equilibrium wages

We will prove the proposition for any given $(q_0, p_0) \in [0,1]^2$. First, let us consider a slightly enlarged parameter space, allowing $\kappa = 0$. Also, let $\Theta = \{\theta \in \mathbb{R}^7_+ : \theta$ is such that Proposition 2 holds$\}$, with typical element $\theta = (\beta, q, \kappa, \overline{w}, r, f(\lambda_m), R_A(r\underline{s}))$. Also, let $\Theta_{-z}$ denote the projection of $\Theta$ in all of the coordinates except for parameter $z$, with typical element $\theta_{-z}$ We will be interested in the projection $\Theta_{-\kappa}$. In order to prove the proposition, we will show that for an open set of parameters, there is always a contract and a corresponding savings level, satisfying the consumer FOC, the zero profit condition, as well as the lower bound for both period wages, such that the individual prefers it to the age contract.

Consider the contract paying the same wage in both periods (and call it the "constant contract"), denoted by $w^*(\theta)$. Let $s(w^*(\theta))$ be the corresponding savings level chosen by the individual, and hence satisfying FOC. Note that both depend on the set of parameters chosen, $\theta$. It can be shown that, for $\kappa \to 0$ there is a unique value $w^*(\theta) > \overline{w} - \kappa$ that implies zero profits when consumers choose their savings level optimally.

Now, let us denote $\triangle U(\kappa, \theta_{-\kappa})$ the difference between the utility level attained with the constant contract and that obtained under the age contract when the individual chooses savings optimally. Note that, since $w^*(\theta)$, and $s(w^*(\theta))$ are continuously differentiable with respect to $\kappa$ for all $\theta_{-\kappa} \in \Theta_{-\kappa}$, $\triangle U(\kappa, \theta_{-\kappa})$ is continuously differentiable in $\kappa$. We will be done if we can show that there exists $0 < \widetilde{\kappa}$ such that, for all $\kappa \in I_{\widetilde{\kappa}} = (0, \widetilde{\kappa})$, $\triangle U(\kappa, \theta_{-\kappa}) > 0$. This would imply that the age contract is not optimal because deviating towards the constant contract would improve individual welfare without violating any of the requirements, and therefore, the age contract would not be chosen for any $\theta \in I = I_{\widetilde{\kappa}} \times \Theta_{-\kappa}$.

First, it is clear that

$$\kappa = 0 \implies \begin{cases} \triangle U(\kappa, \theta_{-\kappa}) = 0 \\ w^*(\theta) = \overline{w} \\ s(w^*(\theta)) = s(\overline{w}, \overline{w}) = s \\ p(w^*(\theta), s(w^*(\theta))) = p(\overline{w}, s) = p \end{cases}.$$

Also, since $s > 0$ solves the consumer problem, it also satisfies the first order condition

$$u'(\overline{w} - s) = p\beta r u'(\overline{w} + rs) + (1 - p)\beta r u'(rs)$$

and this implies that

$$u'(\overline{w} - s) > \beta r u'(\overline{w} + rs).$$

Taking the derivative of $\triangle U(\kappa, \theta_{-\kappa})$ with respect to $\kappa$, using the first order conditions from the consumer problem as well as the implicit differentiation of the equilibrium savings with respect to $\kappa$ and evaluating at $\kappa = 0$ it can be shown that:

$$\frac{\partial \triangle U(\kappa = 0, \theta_{-\kappa})}{\partial \kappa} = u'(\overline{w} - s) > 0 \qquad \forall \theta_{-\kappa} \in \Theta_{-\kappa}.$$

Since $\frac{\partial \triangle U(\kappa=0, \theta_{-\kappa})}{\partial \kappa}$ is continuous, there exists $\widetilde{\kappa} > 0$ and $I_{\widetilde{\kappa}} = (0, \widetilde{\kappa})$ such that:

$$\frac{\partial \triangle U\left(\widetilde{\kappa}, \theta_{-\kappa}\right)}{\partial \kappa} > 0, \qquad \forall \left(\kappa, \theta_{-\kappa}\right) \in I = I_{\widetilde{\kappa}} \times \Theta_{-\kappa}, \text{ with } \mu\left(I\right) > 0.$$

This implies that, by doing a Taylor expansion at $(\kappa = 0, \theta_{-\kappa})$:

$$\forall \left(\kappa, \theta_{-\kappa}\right) \in I, \qquad \triangle U\left(\kappa, \theta_{-\kappa}\right) = \triangle U\left(0, \theta_{-\kappa}\right) + \frac{\partial \triangle U\left(\overline{\kappa}, \theta_{-\kappa}\right)}{\partial \kappa} \kappa > 0$$

for some $\overline{\kappa} \in (0, \kappa)$. This completes the proof.

The economic rationale for this result is that, given $\overline{w}$, and no hiring costs, an increase in the hiring costs will only decrease the present wage in the age contract while diminishing both the present and expected future wage in the constant wage contract. Hence, savings will decrease in the age contract. In the constant contract, however, there are two opposite effects whose total sign is unclear. Thus, given a sufficiently high risk aversion and variance of the density function of the marginal valuation of leisure, savings in the constant wage contract will react less, allowing a smother consumption path.

Analogously, it can be shown that the proposition is non trivial. The economic reason for this result is that, given relatively high hiring costs, if $\overline{w}$ increases starting from zero, the net increase in the second period wage with the age contract more than compensates for the loss of wage in the first period, and, at the same time, increases the probability of working in the second period. Here, the fact that savings cannot react too much because of the non negativity constraint plays a crucial role. Details are available from the authors.

## A.3 Proof of proposition 4. Existence and stability of the steady state equilibrium.

First, note that, given $q_t \in [0, 1]$ proposition 2 ensures that there exists a set of parameter values, dependent on the value of $q_t$ such that there is a unique solution for the problem of generation $t$. Now, we want to strengthen that condition so that the parameter values do not depend on the particular value taken by $q_t$. Since $q_t$ does not appear in the definition of the other parameters, from the proof of proposition 2, such restriction can be accomplished by defining dynamic restrictions as follows. Recall that (13) and (18) define the values of the upper bounds for $f\left(\widehat{\lambda}\right)$. First note that $q_t$ does not appear in the definition of $M_1$. Let:

$$
\begin{aligned}
M_{1,}^* &= \inf_{q_t \in [0,1]} M_1 = M_1 > 0 \\
M_{2,}^* &= \inf_{q_t \in [0,1]} M_2 > 0 \\
M^* &= \min\left[M_1^*, M_2^*\right] > 0.
\end{aligned}
$$

The strict positivity of $M_2^*$ is ensured by the definition of $\pi(q_t, p_t)$ together with the definition of $W_1, W_2$ and $1 - A_1$ (see Appendix A.1). Also in Appendix A.1 the conditions for $g(s_t)$ to be decreasing around $s_t^*$ were established.

Hence, the definition of $M^*$ together with assumptions (4.$i$), (4.$ii$) and (4.$iii$) ensure that, given $q_t \in [0, 1]$, we can apply the proof of proposition 2 to obtain a unique solution to the problem of the generation born at time $t - 1$. Note that part of this solution is $p_t = q_t + (1 - q_t) F\left(\widehat{\lambda}_t\right)$, which allows us to write it as $p_t = p_t(q_t)$. Hence, for generation born at time $t$, it is clear that $q_t$ is the only relevant state variable, and we can write

$$
q_{t+1} = \pi\left(q_t, p_t(q_t)\right) = \pi(q_t)
$$

and note that our previous arguments imply that the solution is differentiable with respect to the state variable $q_t$. Thus, we can obtain an expression for the differential

$$
\frac{dq_{t+1}}{dq_t} = \frac{d\pi(q_t)}{dq_t} = \frac{\partial \pi(q_t, p_t)}{\partial q_t} + \frac{\partial \pi(q_t, p_t)}{\partial p_t} \frac{dp_t}{dq_t}.
$$

If we can prove that, for all $(q_{t-1}, p_{t-1}) \in [0, 1]^2$, $\frac{dp_t}{dq_t}, \frac{\partial \pi(q_t, p_t)}{\partial q_t}, \frac{\partial \pi(q_t, p_t)}{\partial p_t}$ are such that

$$
\sup_{q_t \in [0,1]} \left| \frac{dq_{t+1}}{dq_t} \right| = \sup_{q_t \in [0,1]} \left| \frac{d\pi(q_t)}{dq_t} \right| = \sup_{q_t \in [0,1]} \left| \frac{\partial \pi(q_t, p_t)}{\partial q_t} + \frac{\partial \pi(q_t, p_t)}{\partial p_t} \frac{dp_t}{dq_t} \right| < 1
$$

then $\pi(q_t)$ is a contraction, and by the contraction mapping theorem, it has a unique fixed point $q_t = \pi(q_{t+1}) = q$.[23] Moreover, assumption ($iv$) guarantees that such fixed point is in $(0, 1)$ and we will be done.

---

[23] See Stokey and Lucas (1989), p. 50.

Given $\theta = (q_0, \kappa, f(\lambda_m), \zeta, \overline{w}, \beta, r, R_A(0)) \in \Theta = \{\theta \in \mathbb{R}^8_+\}$, let $q_t = \pi(q_{-1}, p_{-1})$ be known, and $y_t = (w_t, w_{t+1}, s_t, \nu_{1,t}, \nu_{2,t}, \nu_{3,t}, \nu_{4,t})$ be the set of endogenous variables. In order to obtain an expression for $\frac{dp_t}{dq_t}$, let us recall the definition of $p_{t+1}$:

$$p_{t+1} = q_{t+1} + (1 - q_{t+1}) F(u(w_{t+1} + rs_t) - u(rs_t)).$$

Hence,

$$\frac{dp_{t+1}}{dq_{t+1}} = 1 - F\left(\widehat{\lambda}_{t+1}\right) + (1 - q_{t+1}) f\left(\widehat{\lambda}_{t+1}\right) \left[u'(\overline{c}_{t+1}) \frac{dw_{t+1}}{dq_{t+1}} + \frac{\partial \widehat{\lambda}_{t+1}}{\partial s_t} \frac{ds_t}{dq_{t+1}}\right].$$

This value can be shown to be continuous in $q_{t+1} \in [0,1]$. By the extreme value theorem, there exists

$$M_3 = \sup_{q_{t+1} \in [0,1]} \left|\frac{dp_{t+1}}{dq_{t+1}}\right| > 0.$$

Finally, let $M_4 = \max[M_3, 1]$, define

$$\psi = \frac{1}{3M_4},$$

and note that assumption $(iv)$, allows us to place bounds on the derivatives of $\pi$

$$\sup_{q_t \in [0,1]} \left|\frac{\partial \pi(q_t)}{\partial q_t}\right|, \sup_{q_t \in [0,1]} \left|\frac{\partial \pi(q_t)}{\partial p_t}\right| \leq \frac{1}{3M_4}.$$

This implies that

$$\sup_{q_t \in [0,1]} \left|\frac{d\pi(q_t)}{dq_t}\right| \leq \sup_{q_t \in [0,1]} \left|\frac{\partial \pi(q_t)}{\partial q_t}\right| + \sup_{q_t \in [0,1]} \left|\frac{\partial \pi(q_t)}{\partial p_t}\right| \left|\frac{dp_t}{dq_t}\right| \leq \frac{1}{3} + \frac{1}{3M_4} \left|\frac{dp_t}{dq_t}\right| < 1,$$

which completes the proof.

## A.4 Policy assessment: gender based taxation

Let $B$ stand for 'man' and $D$ for 'woman.' Define the 'gross wage gap' as the difference in their expected present value of pre-tax wages and the 'net wage gap' as the difference in the expected present value of their net labor income. Assume that the economy is in a interior stationary steady state characterized by men having a larger marginal product of labor and, as a consequence, a larger unconditional probability of working $\left( \overline{w}^B > \overline{w}^D, \ q^B > q^D \right)$. As mentioned earlier, a technological change that increases the marginal product of women so that it equals that of men would induce a transition phase during which both the expected present value of wages and the unconditional probability of working of women converge to men values. During this transition, the expected present value of wages would be lower for women than for men even though they are equally productive.

Should the government want to speed the transition, it could engage in a budget balanced policy consisting in a lump sum payroll tax on old working men and distributing the proceeds among old working women. Note that, as the tax revenue depends on the market participation rate of old men, the subsidy each women receives is also dependent on it. Therefore, such a tax scheme links two markets that initially were independent of each other, or more precisely, makes the outcomes in the market for female workers to be dependent on the equilibrium values in the market for male workers. It should also be underscored that, as this policy modifies the stationary value of $q^j$, it needs to be dynamically adjusted, and eventually eliminated. Otherwise it would lead to a new stationary steady state in which the unconditional probability of working of women was larger than that of men. An indicator of the gap between the two groups is simply $q_t^B - q_t^D$. We argue that this difference decreases faster with the aforementioned policy.

As we are assuming that the government maintains its budget balanced at every time $t$, the amount of the subsidy an individual receives is determined at each time by the tax revenue and the participation rates of both groups. Let $T^t = \{T_{t'}\}_0^\infty$ denote a tax scheme where $T_{t'+1}$ is the tax faced by the generation born at time $t'$, for all $t' \leq t$ but is 0 for all generations born at time $t' > t$. Let $q_{t+1}^j \left( T^t \right)$ denote the unconditional probability of working of type $j$ individuals born at time $t$ given that the previous generations have paid the taxes implied by tax scheme $T^t$. Recall that $q_t^B \left( 0 \right) > q_t^D \left( 0 \right), \forall t \in \mathbb{N}$, i.e., there exists a positive gap between the unconditional probability of working of both groups when there is no tax. Finally, let $q^j \left( T \right)$ denote the unique steady state value of type $j$ individuals unconditional probability of working if every generation had faced a fixed tax $T$.

**Proposition 5** If:

(5.$i$) the conditions for the existence of a steady state (Proposition 4) hold

(5.$ii$) $\kappa$ is such that the equilibrium wages depend on the unconditional probability of working

for all $t$

then there exists $\overline{T} > 0$ such that a budget balanced government policy consisting of a flat rate tax scheme $(T^t)$ on type $B$ working old and a subsidy of $\frac{T_{t+1}w_{t+1}^B P_{t+1}^B}{P_{t+1}^D}$ to type $D$ working old, will cause both the lifetime gross and net wage gap to shrink. In addition, unconditional probabilities of working would converge faster in the sense that

$$\left. \frac{\partial \left( q_{t+1}^B \left( T^t \right) - q_{t+1}^D \left( T^t \right) \right)}{\partial T^t} \right|_{T^t=0} < 0,$$

for all $t$ and $0 < T_t < \overline{T}$

**Proof** First, let $\kappa$ be small but strictly positive. In order to prove the claim, we first construct the problem of type $B$ and $D$ individuals, as well as the problem of the firm. Since we impose a payroll tax $T_{t+1}$ on type-$B$ individuals that decide to work when old, and use the revenues to subsidize type-$D$ individuals that work when old, the time $t$ constraint is the same as in the original problem for both individuals while time $t + 1$ constraints are now given by $c_{t+1}^B = s_t^B r + \left( w_{t+1}^B - T_{t+1} \right) L_{t+1}^B$ and $c_{t+1}^D = s_t^D r + \left( w_{t+1}^D + T_{t+1} \frac{P_{t+1}^B}{P_{t+1}^D} \right) L_{t+1}^D$ where $P_{t+1}^j = E_j \left[ p_{t+1}^j \right]$ stands for the expectation across type $j$ agents born at time $t$. Hence, the previous definitions of $c_t^j, \overline{c}_t^j, \underline{c}_t^j$ are modified accordingly. When making their decisions, both individual agents and firms take $T_{t+1}$ and $q_{t+1}^j$ as given with respect to their actions.

As in Appendix A.1, let us define the problem of the firm and note that in this case, firms face an additional constraint, given when facing type $B$ individuals:

$$w_{t+1}^B + s_t^B r - T_{t+1} \geq 0 \quad \left( \nu_{5,t}^B \right).$$

To ensure that under our assumptions, the new problem of the firm is well defined, and has a unique solution, we restrict $T_t \in [0, \overline{w} - \kappa]$, for all $t$, so that the new constraint is never binding and we can eliminate it from the problem. Denote by $V^j \left( s_t^j, q_{t+1}^j, T^{t+1} \right)$ the problem of the firm when facing individual of type $j$ and by $\Omega^j \left( s_t^j, q_{t+1}^j, T^{t+1} \right)$ its constraint set, $j = B, D$ As in Appendix A.1 and A.2, our assumptions ensure that the firm problem is well defined and has a solution. Moreover, by a reasoning analogous to that of previous appendices, the individual utility is strictly quasiconcave in wages, it has non-zero gradient, and the problem satisfies the constraint qualification at any solution. Hence, for a given triplet $\left( s_t^j, q_{t+1}^j, T^{t+1} \right)$, the unique solution is totally characterized by the Kuhn-Tucker conditions. On the other hand, our assumptions also guarantee that given $\left( T^{t+1}, q_{t+1}^j \right)$ the FOC of the individual agent problem has a unique solution. Note that a change in $T^{t+1}$ will have two effects: first it will affect the transition for each generation $t$ and, second it will affect the steady state value of both quit rates. Define now

the lifetime gross wage gap,

$$w_t^B + w_{t+1}\frac{p_{t+1}^B}{r} - w_t^D - w_{t+1}^D\frac{p_{t+1}^D}{r} = \overline{w}\left(\frac{p_{t+1}^B}{r} - \frac{p_{t+1}^D}{r}\right),$$

by the zero profits condition. Also, recall that $q^j(T)$ denote the stationary value of $q^j$ when all generations face a tax of amount $T$.

By differentiating the system of equations that defines the equilibrium for a given generation, and evaluating at the equilibrium, it can be shown that the following conditions hold

$$i) \quad \left.\frac{dp_{t+1}^B}{dT^{t+1}}\right|_{T^{t+1}=\{0\}_0^\infty} < 0, \qquad \left.\frac{dp_{t+1}^D}{dT^{t+1}}\right|_{T^{t+1}=\{0\}_0^\infty} > 0$$

$$ii) \quad \left.\frac{dq_{t+2}^B}{dT^{t+1}}\right|_{T^{t+1}=\{0\}_0^\infty} < 0, \qquad \left.\frac{dq_{t+2}^D}{dT^{t+1}}\right|_{T^{t+1}=\{0\}_0^\infty} > 0$$

$$iii) \quad \left.\frac{dp_{t+1}^j}{dq_{t+1}^j}\right|_{T^{t+1}=\{0\}_0^\infty} > 0$$

$$iv) \quad \left.\frac{dq^B(T)}{dT}\right|_{T=0} < 0, \qquad \left.\frac{dq^D(T)}{dT}\right|_{T=0} > 0.$$

Taken together, these inequalities imply the result.

Although the intuition behind this result is straightforward, it is worth noting that the proof is not trivial, because of the endogenous nature of contracts. Take the contract of men: anticipating the tax men will face when old, firms will adjust the wage scheme offered to diminish its impact on the worker welfare. In particular firms will offer a lower wage to young men and a higher one to old workers, because this lessens the impact of the tax, as it shifts income from the 'low' to the 'high' marginal valuation of consumption period of the life of the worker. This could lead men to increase their likelihood of working when old. Proposition 5 shows that this is not the case, and men end up with a reduced probability of working.