Feeling guilty and redistributive politics

Gilles Le Garrec

To cite this version:

Gilles Le Garrec. Feeling guilty and redistributive politics. 2009. <hal-01066215>

HAL Id: hal-01066215

https://hal-sciencespo.archives-ouvertes.fr/hal-01066215

Submitted on 19 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Centre de recherche en économie de SCIENCES PO

Document de travail

« Feeling guilty and redistributive politics »

N° 2009-22
September 2009

Gilles Le Garrec
(OFCE-Sciences-Po)

Observatoire Français des Conjonctures Économiques
69, Quai d'Orsay 75340 Paris Cedex 07
Tel : 01 44 18 54 00 Fax : 01 45 56 06 15
E-mail: ofce@ofce.sciences-po.fr Web: http://www.ofce.sciences-po.fr

Abstract

In mainstream economics individuals are supposed driven only by their self interest. By contrast, in this article, in the line of a "new synthesis" in moral psychology we assert that the voting behavior over redistribution is best characterized by first an automatic cognitive process which generates quicly intuitions on the fair level of redistribution, and second by a rational self oriented reasoning which controls the feeling of guilt associated with the fair intuitions. As a result of this dual-process decision-making, we show that the U-shape between inequality and redistribution supported by the data is a general feature of the model. In addition, assuming that the feeling of guilt is context dependent and is reduced if the previous generation failed in implementing the intuitively fair institution, the model exhibits a multiplicity of steady states which can explain the huge difference of redistribution observed between Europe and the United States.

Keywords: redistribution, voting behavior, fairness, psychology

JEL: H53, D72, D64, D87

1 Introduction

In mainstream economics individuals are supposed driven only by their self interest. As a consequence, when studying the redistributive phenomenon in democracy, the first challenge for most economists is to explain why there is so little redistribution in democracy. Indeed, considering that the idea of democracy is captured by the majority rule, as the median citizen is characterized by an income lower than the average, a majority should support a complete redistribution of income to satisfy his self interest. As a canonical answer to this issue, Meltzer and Richard (1981) have showed that selfish people have no interest to support a too high redistribution, even if they are poorer than average, because of a tax disincentive effect that lowers productivity. Their model also implies that we should observe a growing relationship between redistribution and income inequality. However, such relationship is weakly supported by data. While redistribution is higher in Europe than in the United States, their pre-tax income inequalities appear similar (see Table 1). By contrast, Perotti (1996), Moene and Wallerstein (2001) and Iversen and Soskice (2006) support that the empirical relationship between inequality and redistribution is the opposite of the predicted one; or it exhibits a U-shape as suggested by de Mello and Tiongson (2006).

Countries	Pretax-tax income inequality (GINI)	Public Social Spendings (% GDP)
France	0.48	29.2
Germany	0.51	26.7
Sweden	0.43	29.4
UK	0.46	21.3
US	0.46	15.9

Table 1. Income inequality and social spending in 2005 (source: OECD).

In order to improve the canonical model's predictions, four main directions, economic, political, institutional and behavioral, have been investigated[1]. From an economic perspective, considering imperfections in the credit and in the insurance markets, Bénabou (2000) has stressed an enhancing productivity effect of redistribution which can balance the tax discentive effect and lead to a multiplicity of equilibria. From a political perspective, it has been argued that rich people have more influence in the political process because of their higher propension to vote, their campaign contributions or their lobbying activities, and then that the pivotal voter is not characterized by the median income (see Rodriguez, 2004, Campante, 2006, Petrova, 2008). In addition, Roemer (1998) supports that political choices are in essence mul-

[1]See Alesina and Angeletos (2004), Lind (2005) and Campante (2006) for an overview.

tidimentionals and therefore that stable coalitions cannot be characterized only by the income of their members. From an institutional perspective, Iversen and Soskice (2006) have argued that the two-party majoritarian system which characterizes the United States lowers the pressures for redistribution compared to a multiparty proportional system as in Europe. At least, from a behavioral perspective, Bénabou and Ok (2001) and Bénabou and Tirole (2006) have pointed out that poor people do not support high redistribution because they expect becoming rich, eventually by manipulating their perception of the social mobility. The postulate that individuals are driven only by their self interest has also been challenged when studying redistribution (Piketty, 1995, Alesina and Angeletos, 2005, Lind, 2007).

It has been challenged first because an impressive number of experimental studies have pointed out that individuals do not behave selfishly in the way supposed in mainstream economics (see Batson, 1991, Fehr and Schmidt, 2006). Evidence appears particularly clear in the Ultimatum Game (Güth et al, 1982). In this game, two anonimous subjects must agree in the split of a given amount of money ($10). One participant, the proposer, can make one proposal on how to split the money. The other one, the recipient, can either accept or reject the proposal. If he accepts, the proposal is carried out, if not both players get nothing. In theory, in such a one-shot game the proposer should offer to the recipient an amount of money as low as possible, and the recipient should accept any proposal superior to 0. By contrast, robust findings in the lab show an average proposal of \$4 with lots of proposed 50/50 splits. Besides, any proposal below \$2.50 has a high probability to be rejected and about 25% of proposals are rejected. The Ultimatum Game stresses behaviors which are characterized by fairness and inequity aversion.

The fact that people can act fairly as underlined in experimental studies does not entail that the standard assumption of human selfishness is not relevant in most domain of economics. As noted for example by Adam Smith (in *The Wealth of Nations*, 1776), "*It is not from the benevolence of the butcher, the brewer, or the baker, that we can expect our dinner, but from their regard to their own interest.*" But studies from Fong (2001), Corneo and Grüner, (2002), Alesina and La Ferrara, (2005), Corneo and Fong, (2008) and Alesina and Giuliano (2009) clearly show that people do care about fairness in their demand for redistibution. In addition, for Alesina, Glaeser and Sacerdote (2001) the fact that beliefs according to which *luck rather than effort determines income*[2] are strong predictors of the level of redistribution (as suggested in Figure 1) is another evidence supporting that fairness, unlike income inequality, has a major influence in shaping redistributive politics.

In the utilitarian trend, the notion of fairness in economics appears log-

[2]From World Values Survey data, they highlight that 54% of Europeans versus 30% of Americans believe that *luck rather than effort determines income.*

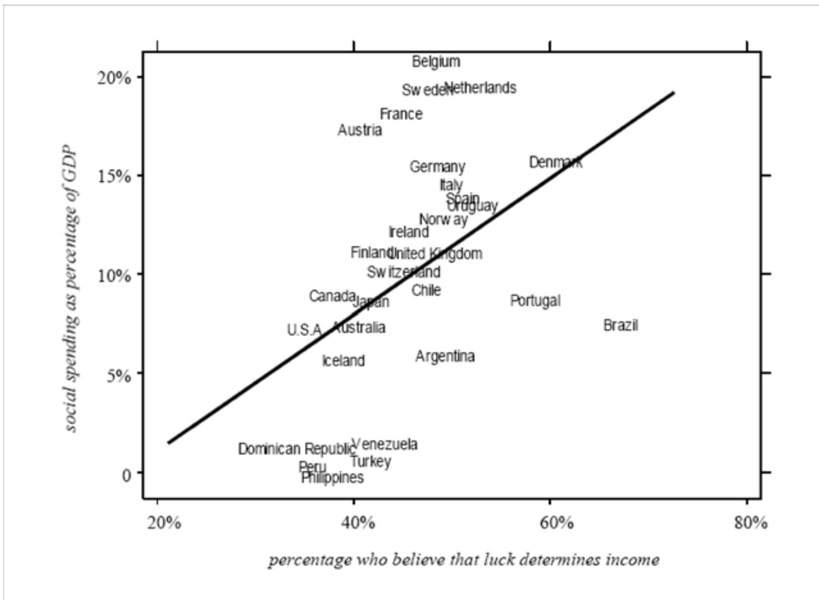

Figure 1: Social spending and luck (source: Alesina, Glaeser, Sacerdote, 2001).

ically associated with a function which aggregates individual utilities, the social welfare. In this line, if voters care only about the welfare of the whole population when considering redistribution, Piketty (1995) has showed that international differences in the level of redistribution when countries share identical economic fundamentals can be explained by different beliefs about social mobility sustained by an imperfect learning process. Closer from a concept of reciprocal altruism, Lind (2007) considers the case where voters care about their self interest and the welfare of the members of their own group more than the rest of the population. In such a context, he shows that both fractionalization and group antagonism reduce redistribution. In the line of Gilens (1999), he supports then that the difference of redistribution between Europe and the United States can be explained by a difference of ethnic fractionalization. In Alesina and Angeletos (2005), voters also care about both their self interest and fairness. In their model, fairness is not defined according to an utilitarian social welfare as in Piketty (1995) and Lind (2007), but according to the deontological principle that *everyone should receive what he deserves* whose relevance is empirically supported in psychology and sociology (see Schokkaert, 1998, Forsé and Parodi, 2006). With income depending on effort and luck, they show that cultural variability arises as a multiplicity of equilibria resulting from different self-fulfilled beliefs. By expecting low redistribution, Americans invest in their human capital and generate conditions for low redistribution by reducing the importance of luck in the income determination. Conversely, by expecting a high redistribution, Europeans invest less in their human capital and will support later a high redistribution.

In this article, we consider that voters are influenced both by fairness (in the way of Alesina and Angeletos, 2005) and their self interest conserning redistribution. In addition, in the line of recent findings in neurosciences and cognitive psychology, we also explicitly consider that fair and selfish motives are generated by two distinct cognitive processes whose different features (see Appendix A) are of interest in explaining cultural variability of redistributive politics. First, as argued by Haidt (2001, 2007, 2008; see also Nado et al., 2006), an automatic and domain-specific cognitive process[3][4] quickly generates and relates moral intuitions (shaped according to Hauser, 2006, and Mikhail, 2007, by a Universal Moral Grammar or UMG) then emotions such as guilt. Supporting this thesis, a large number of both neuroimaging and neuropsychological studies (Greene et al., 2001, 2004, Berthoz et al., 2002, 2006, Moll et al., 2002, 2005, Decety and Chaminade, 2003, Hsu et al., 2008; Damasio, 1994, Anderson et al., 1999, Blair, 2001, Koenigs and Tranel, 2007; see Greene, 2005, Young and Koenigs, 2007 for reviews) show that fair behaviors are associated with brain areas involved in emotional processing (amygdala, insula, ...). For example, considering the Ultimatum Game, Sanfey et al. (2003) have showed that the rejection of an unfair proposal (assimilated to a fair punishment) was related to the activation of the anterior insula. Besides, the rejection of an unfair proposal has also been associated with an increased skin conductance which reveals the emotional intensity of the choice (van 't Wout et al., 2006). Emotions are then integrated into a self oriented rational cognitive process (system 2; see Appendix

[3] As explained by Fehr and Schmidt (2006), "*the term automatic in this case refers to a process that does not require conscious and effortfull processing but which can nevertheless be inhibited or controlled.*"

[4] Such a cognitive process is named system 1 by Kanhneman, 2003; see also Camerer et al. 2005, Evans, 2008, and Appendix

A) such that the human behavior is equally driven by emotions and by the rational reasoning standard in economics (Gray, 2004, Bechara, 2004, Cohen, 2005, Pessoa, 2008).

The rest of the article is organized as follows. In section 2, we propose a model in the line of the dual-process theory of decision-making. An intuitively fair level of redistribution is generated by an automatic cognitive process (UMG), and the related feeling of guilt is controlled by a rational self oriented reasoning as described in Figure 1. We then show that the U-shape between inequality and redistribution which is supported by the data is a general feature of the model. In section 3, assuming that the feeling of guilt is context dependent and is reduced if the previous adult generation failed in implementing the intuitively fair level of redistribution, we show that the model exhibits a multiplicity of steady states which can explain the huge difference of redistribution observed between Europe and the United States. In section 4, we generalize the multiplicity of steady states result when taking into account family history and the intergenerational transmission of inequality. We conclude briefly in a last section.

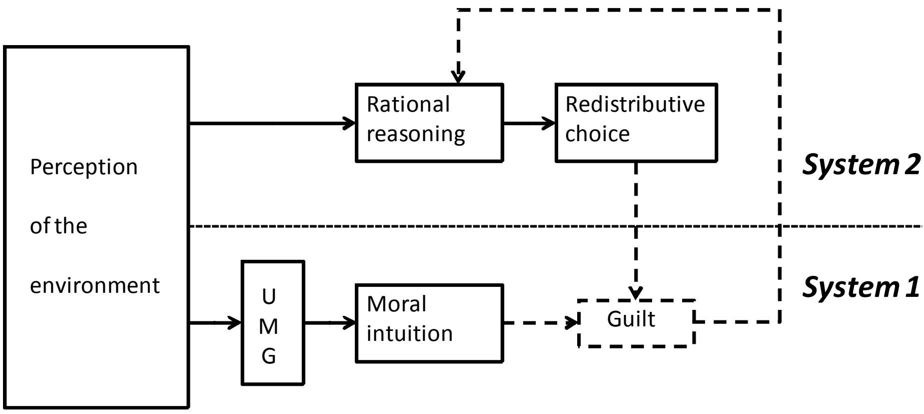

Figure 2: Dual-process decision-making and voting over redistribution

2 Redistribution: from moral intuition to the rational control of guilt

The economy is populated by a continuum of mass one of individuals whose actions take place according to the timeline on Figure 2. For each individual of type i, the second period pre-tax income is

$$y_i = a_i + e_i + \varepsilon_i \tag{1}$$

where e_i is the level of effort he makes in period 1, a_i is a positive parameter which represents his intrinsic motivations[5] (known in period 1), and ε_i

[5]Intrinsic motivations are motivations which are unrelated (or eventually negatively related) to external rewards such as income or status (see Ryan and Deci, 2000, Bénabou and Tirole, 2002).

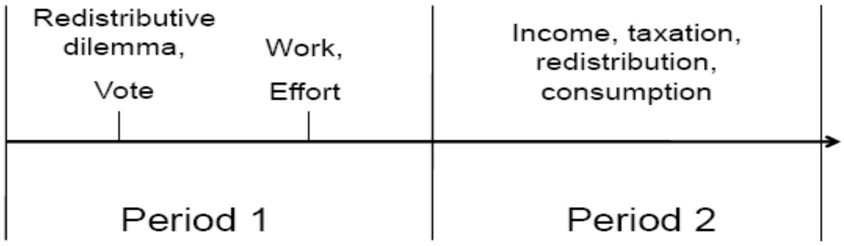

Figure 3: Timing of actions

represents his luck (or bad luck) in period 2, unknown in period 1 such that $E_1[\varepsilon_i] = 0$ and independent from his intrinsic skills.

Redistribution is characterized by a taxation rate τ and a lump sum allocation g. If we suppose a balanced budget, then $g = \tau\bar{w}$.

2.1 Effort

Let us assume that the rational reasoning is characterized by the maximization of the consumption in period 2 while effort involves a utility loss equal to $\frac{e_i^2}{2\alpha_i}$ in period 1, where α_i represents a taste for effort. The rational reasoning (system 2) of an individual i can therefore be characterized by the following function:

$$U_i = y_i (1 - \tau) + \tau\bar{y} - \frac{e_i^2}{2\beta_i \alpha_i} \tag{2}$$

where $\beta_i < 1$ represents procrastination or willpower (see Bénabou and Tirole, 2006). If the choice of effort entails no moral conflict, the optimal effort for each individual corresponds to the maximization of $E_1[U_i]$ and is characterized by:

$$e_i = \beta_i \alpha_i (1 - \tau) \tag{3}$$

As usual, we can observe that redistribution reduces the effort incentive. Assume then that cognitive skills behind intrinsic motivations are perfectly correlated with taste for effort and willpower such that $a_i = \beta_i \alpha_i$. With (3), the pre-tax income (1) can be rewritten as:

$$y_i = a_i (2 - \tau) + \varepsilon_i \tag{4}$$

As the level of effort is reduced by redistribution, of course the pre-tax is reduced. As a consequence, redistribution will reduce not only the variance of the disposable income, but also the variance of the pre-tax income.

2.2 Moral intuition

In the line of the dual-process theory of decision making (Kaneman, 2003) and the "new synthesis" in moral psychology (Haidt, 2001, 2007, 2008; see also Nado et al., 2006), let us assume a cognitive process which automatically and quickly generates moral intuitions. Assume in addition that these intuitions when considering income distribution are characterized by the deontological principle that *each people should receive what he deserves*[6], where the deserved or fair income is defined by:

$$\hat{y}_i = a_i + e_i \tag{5}$$

i.e. the income without luck and bad luck[7]. Following Alesina and Angeletos (2005), we then characterize the universal sense of fairness by:

$$F = \int_i \left\{ \left[(1-\tau)\, y_i + \tau \bar{y} \right] - \hat{y}_i \right\}^2 di \tag{6}$$

The fair-motivated cognitive process characterized by (6) urges people to reduce unfairness. As α and ε are independently distributed, (4) allows to rewrite (6) as (see Appendix B):

[6] Forsé et Parodi (2006) show that European countries share an identical hierarchy of moral principles: first the guarantee of basic needs, second fairness (merit), and far less important equality of income. If we admit that basic needs are mostly satisfied in Europe and in the United States, fairness is the relevant concept to study marginal variations of the redistribution levels.

[7] The existence of a universal fair income distribution supposes that it must be shared by each individual independently of his own income. On French data, Piketty (2003) shows that such an assumption is supported.

$$\frac{F}{\sigma_a^2} = (1-\tau)^2 L + \tau^2 (2-\tau)^2 \tag{7}$$

where $L = \frac{\sigma_\varepsilon^2}{\sigma_a^2}$, σ_ε^2 being the variance of ε which represents the importance of luck in the income determination (one can think of a simple case where $\varepsilon = \pm l$ with probability $\frac{1}{2}$, and then where $\sigma_\varepsilon^2 = \frac{1}{2}l^2$) and σ_a^2 being the variance of a. L represents then the relative importance of luck in the income determination and the *intuitive* fair tax rate which corresponds to the minimization of (7) is as follows:

$$\tau^f = \begin{cases} 1 - \sqrt{1 - \frac{L}{2}} & \text{if } L \leq 2 \\ 1 & \text{otherwise} \end{cases} \tag{8}$$

Under this fair motive, the tax rate is only increasing with the relative importance of luck in the income determination: $\frac{\partial \tau^f}{\partial L} > 0$. As the relative importance of luck is defined by $L = \frac{\sigma_\varepsilon^2}{\sigma_\alpha^2}$, it also means that an increase of the variance of a reduces the relative luck and then $\frac{\partial \tau^f}{\partial \sigma_\alpha^2} < 0$. In addition, let us admit that an increase of the variance is correlated with an increase between mean and median incomes. For example, assume that a is equal to $a_0 = 1$ with probability $\rho > \frac{1}{2}$ and $a_1 > 1$ with probability $1-\rho$. In such a case, setting $\bar{a} - a_{med} = \Delta$, we can show with simplicity that $\sigma_a^2 = \frac{\rho}{1-\rho}\Delta^2$ and then that $\frac{\partial \sigma_a^2}{\partial \Delta} > 0$ (similar results could be obtained using standard statistical distributions such as Pareto and Lognormal). It follows that an increase of the difference between mean and median incomes is associated with a lower intuitive fair level of redistribution.

2.3 Guilt, rational control and the demand for redistribution

For any level of redistribution that an individual could choose, compared to his intuition of what is fair, he will have a feeling of guilt. We can reasonably assume that this feeling of guilt will be all the stronger as the potential redistributive choice τ will be far from the fair level τ^f. Let us then define the self-conscious experience of guilt as follows:

$$\frac{\varphi}{2}\left(\tau^f - \tau\right)^2 \tag{9}$$

If an individual is only driven by his emotion of guilt, he will always choose the fair level of redistribution. But, as explained by Gray (2004), "*At some point of processing, functional specialization is lost, and emotion and cognition conjointly and equally contribute to the control of thought and behavior*", i.e. emotion is integrated into the rational reasoning or is controlled by this rational process. The voting behavior in period 1 is therefore characterized by maximizing:

$$E_1\left[y_i\left(1-\tau\right) + \tau\bar{y} - \frac{e_i^2}{2\beta_i}\right] - \frac{\varphi}{2}\left(\tau^f - \tau\right)^2 \tag{10}$$

Considering (4), it follows that the optimal level of redistribution for an individual of type i is:

$$\tau_i = \begin{cases} \frac{2(\bar{a}-a_i)+\varphi\tau^f}{2\bar{a}-a_i+\varphi} & \text{if } a_i \leq \bar{a} + \frac{\varphi\tau^f}{2} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Knowing that τ_i (defined by 11) is a decreasing function of a_i and assuming that the distribution of a is skewed to the right, i.e.$a_{med} \leq \bar{a}$, it entails that the tax rate chosen under the majority rule is characterized by:

$$\tau^* = \frac{2\left(\bar{a} - a_{med}\right) + \varphi\tau^f}{2\bar{a} - a_{med} + \varphi} \tag{12}$$

Let $\tau^s = \frac{2(\bar{a}-a_{med})}{2\bar{a}-a_{med}}$ be the taxation rate if people were driven only by their self interest, i.e. if $\varphi = 0$, we can rewrite (12) as:

$$\tau^* = \xi\tau^s + (1-\xi)\,\tau^f \tag{13}$$

where $\xi = \frac{2\bar{a}-a_{med}}{2\bar{a}-a_{med}+\varphi}$. As $a_{med} \leq \bar{a}$, $0 < \xi \leq 1$ and the level of taxation τ^* is then a convex combination of the *selfish* level τ^s and the fair one τ^f, where the weighting depends only of the distribution of a. It follows that:

Proposition 1 *Any increase of the importance of luck in the income determination, everything else being equal, results in the increase of the level of redistribution.*

As the variance of ε only affects the relative importance of luck in the wage determination L, proposition 1 is straightforward: any increase of σ_ε^2 increases L, then τ^f and τ^*.

Proposition 2 *Assume that* $\frac{\partial \sigma_a^2}{\partial \Delta} > 0$, $\forall 0 < \sigma_a^2 \leq \frac{\sigma_\varepsilon^2}{2}$ $(L \leq 2)$ *and* $\forall 0 < \varphi < \infty$*, the level of redistribution exhibits a U-shape with the difference between mean and median incomes (see Appendix C).*

The impact of an increase of the difference between mean and median incomes appears complex. First, as usual in the standard model, such an increase leads to an increase of the *selfish* tax rate. For example, assume again that a is equal to $a_0 = 1$ with probability $\rho > \frac{1}{2}$ and $a_1 > 1$ with probability $1 - \rho$. With $\bar{a} - a_{med} = \Delta$, it follows that $\tau^s = \frac{2\Delta}{1+2\Delta}$ and then that $\frac{\partial \tau^s}{\partial \Delta} > 0$. But at the same time $\sigma_a^2 = \frac{\rho}{1-\rho}\Delta^2$ and then $\frac{\partial \sigma_a^2}{\partial \Delta} > 0$. An increase of the difference between mean and median incomes is then also related to an increase of σ_a^2 which reduces the relative importance of luck in the income determination and therefore τ^f (eq. 8). The total effect on τ^* is then uncertain and depends of the weighting ξ. We can thereafter observe that ξ is an incresing function of the difference between the mean and the median incomes. In other words, the importance of τ^s (respectively τ^f) in the determination of τ^* is all the more (all the less) important that the difference between mean and median incomes is strong. It follows that the leading effect when Δ is small is supported by τ^f and in this case any increase of Δ (which is associated with an increase of σ_a^2 and then a decrease of L) would result in a decrease of τ^* whereas the leading effect when Δ is strong is supported by τ^s and any increase of σ_a^2 in this case would result in an increase of τ^*. The U-shape between inequality and redistribution which is supported by the data is then a general feature of the model. Nevertheless, at this point, we can not assert that this result is specifically due to the distinction between system 1 and system 2. Indeed, in a more standard approach where fairness would be part of a global reasoning, i.e. part of a stable utility function as in Alesina and Angeletos (2005), a small level of redistribution considering the fair concern would generate a large desutility which is similar to our concept of guilt. Therefore, as long as we consider that the difference of the level of redistribution observed between Europe and the United States is only sustained by different exogenous parameters (σ_a^2 and σ_ε^2), we can deduce that the existence of a universal sense of fairness helps understanding the U-shape between inequality and redistribution whatever the distinction between system 1 and system 2.

Consider alternatively that the cultural variability could persist even with identical economic fundamentals. In that case, emphasizing on the context-dependency of the feeling of guilt (Markus and Kitayama, 1991, Scherer, 1997, Tracy and Robins, 2004) appears particularly relevant. Indeed, if we follow the cultural trend in psychology, the feeling of guilt can both create and be shaped by cultural practices and institutions, and the cultural variability of the redistributive intitutions may arise from multiple steady states.

3 Feeling guilty and context dependency

In the cognitive trend of psychology, in the line of evolutionary theory, emotion is often viewed as a universal set of largely prewired internal processes of self-maintenance and self-regulation (LeDoux, 1996, Ekman, 1992). Under this biological aspect of emotion, nonverbal measures such as facial expressions, skin conductance or the activation of certain parts of the brain (amygdala, insula, ...) have been favored to assess the universality of the emotional phenomena (see Phelps, 2009). Using facial expressions, Ekman and Friesen (1971) have for example suggested that there exist six basic emotional expressions (happy, sad, fear, anger, disgust and surprise) whose characteristics are universally observed across culture. But, in the cultural trend of psychology closer to anthropology, along with these six basic emotions exists self-conscious emotions such as guilt or pride which are complex social emotions whose experience is not characterized by universality (Markus and Kitayama, 1991, Mesquita and Frijda, 1992, Frijda and Mesquita, 1994, Scherer, 1997, Eisenberg, 2000, Tracy and Robins, 2004, Goetz and Keltner, 2007, Edelstein and Shaver, 2007). To illustrate their point, Tracy and Robins (2004) explain that a person may feel great happiness after winning either a lottery or an athletic event, but that only the athletic success can generate pride. Following this reasoning, we argue that if an athletic success can generate pride, it can also generate guilt if there is cheating or drug use, but obviously the self-conscious exprience of guilt will be the less important that the drug use is generalized and established as a norm. In other words, a self-conscious emotion is context-dependent because it is based on a self evaluation process which requires comparison with others' behaviors and cultural practices. As an evidence, Harlé and Sanfey (2007) and Twenge et al. (2007) have showed that manipulating the environment in prosocial behavior experiments (Ultimatum Game in Harlé and Sanfey, 2007) can change significantly the results. In addition, prosocial behavior experiments (among which UG) undertaken in small-scale societies suggest that "*culturally transmitted behavioral variation may substantially affect decision-making*" (Henrich, 2000, Henrich et al., 2001, 2005).

To take into account the cultural side of human behaviors, consider now an OLG model where individuals live for two periods, childhood and adulthood. During the first period, child are socialized. When adult, individuals vote, work and consume according to the same timeline already described in Fig. 1 (see Fig.). In addition, let us assume now that, even if we still

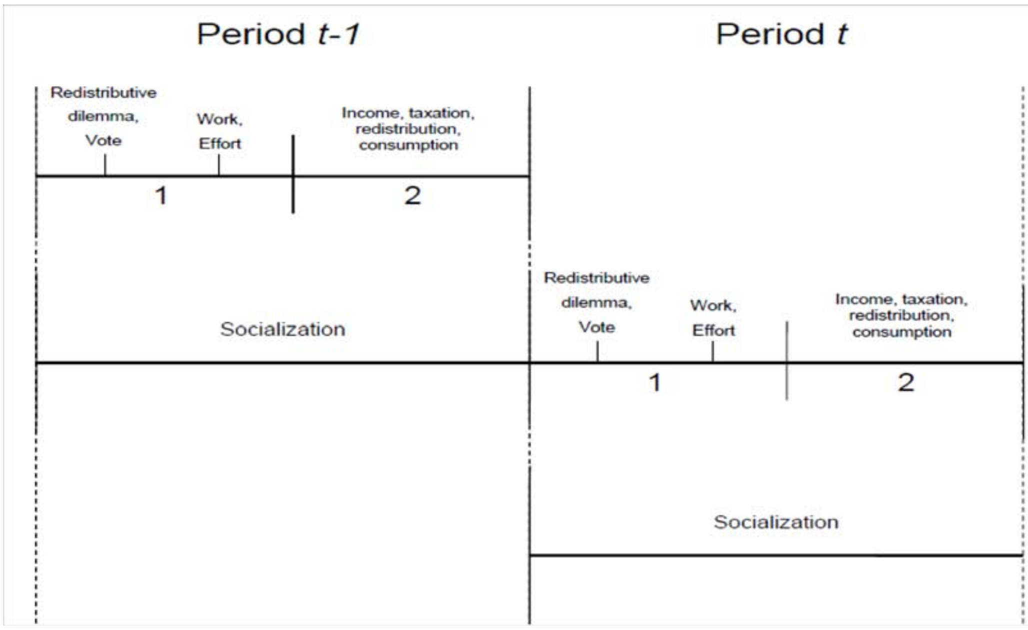

Figure 4: Timing of actions when considering socialization

consider that the moral intuition is universal, the self-conscious expression of guilt is shaped by culturally transmitted (during childhood) practices and institutions. More precisely, let us assume that unfairness of the institution observed by a person during his childhood, chosen by the previous adult generation, reduces his conscious feeling of guilt when adult, we can represent the self-conscious experience of guilt by:

$$\frac{\Phi}{2\Gamma_{t-1}}\left(\tau^{f}-\tau\right)^{2} \tag{14}$$

where $\Gamma_{t-1}=\left[\tau_{t-1}^{f}-\tau_{t-1}^{*}\right]^{2}$ represents the unfairness of the institution chosen by the previous adult generation in $t-1$.

In such a case, by adapting directly (13), we deduce that the level of redistribution in t is characterized by:

$$\tau_t^* = \xi_t \tau^s + (1 - \xi_t)\,\tau^f \tag{15}$$

where

$$\xi_t = \frac{2\bar{a} - a_{med}}{2\bar{a} - a_{med} + \frac{\Phi}{\left[\tau_{t-1}^f - \tau_{t-1}^*\right]^2}} \tag{16}$$

It follows that:

Proposition 3 *If* $\tau^f - \tau^s > 2\sqrt{\frac{\Phi}{2\bar{a} - a_{med}}}$, *the model exhibits two steady states characterized by:*

$$\tau^* = \begin{cases} \tau^f \\ \frac{1}{2}\left(\tau^f + \tau^s - \sqrt{\left(\tau^f - \tau^s\right)^2 - \frac{4\Phi}{2\bar{a} - a_{med}}}\right) \end{cases}$$

where $\tau^f = \begin{cases} 1 - \sqrt{1 - \frac{L}{2}} & \textit{if } L \leq 2 \\ 1 & \textit{otherwise} \end{cases}$ *and* $\tau^s = \frac{2(\bar{a} - a_{med})}{2\bar{a} - a_{med}}$.

The multiplicity of steady states results from the interaction of the two distinct processes. First, guilt that people can feel in relation with their moral duty shapes the voting behavior: the stronger is this feeling, the closer from fair intuition is their choice. Second, if people are socialized in a cultural environment whose cultural practices and institutions do not reflect the intuitive fairness, internalization of the observed norm *it is allowed to* *act unfairly*, or *you should behave according to your own interest*, will reduce individual responsability and therefore the feeling of guilt considering the moral duty. By contrast, if people are socialized in a cultural environment whose cultural practices and institutions reflect the intuitive fairness, the norm *it is not allowed to act unfairly* observed reinforce the feeling of guilt. We therefore observe, if $\tau^f - \tau^s > 2\sqrt{\frac{\Phi}{2\bar{a}-a_{med}}}$, that the convergence towards the two different steady states depends only from the initial level of taxation as stated in proposition 2.

Proposition 4 *Consider the tax sequence* $\{\tau_t\}_{t=0}^{t=\infty}$ *and suppose that* $\tau^f - \tau^s > 2\sqrt{\frac{\Phi}{2\bar{a}-a_{med}}}$, *if* $\tau_0 \in \left]\tau^f - \delta; \tau^f + \delta\right[$, *where* $\delta = \frac{\tau^f - \tau^s - \sqrt{\left(\tau^f - \tau^s\right)^2 - \frac{4\Phi}{2\bar{a}-a_{med}}}}{2}$, *then* $\lim\limits_{t\to\infty} \tau_t = \tau^f$, *while if* $\tau_0 \notin \left]\tau^f - \delta; \tau^f + \delta\right[$ $\lim\limits_{t\to\infty} \tau_t = \frac{1}{2}\left(\tau^f + \tau^s - \sqrt{\left(\tau^f - \tau^s\right)^2 - \frac{4\Phi}{2\bar{a}-a_{med}}}\right)$.

If the initial level of taxation corresponds to an institution sufficiently closed to the fair level, the feeling of guilt will be reinforced and people next period will vote for an even closer from fair level of redistribution. This *emotional contagion* process ends with the implementation of the fair institution, $\tau^* = \tau^f$. By contrast, starting from a point where the institution is sufficiently far from the fair level leads to a too strong signal *you should behave according to your own interest* which reduces concerns about morality and then the associated feeling of guilt, preventing from an emotional contagion to the benefit of a rational selfish control. The process will end at a

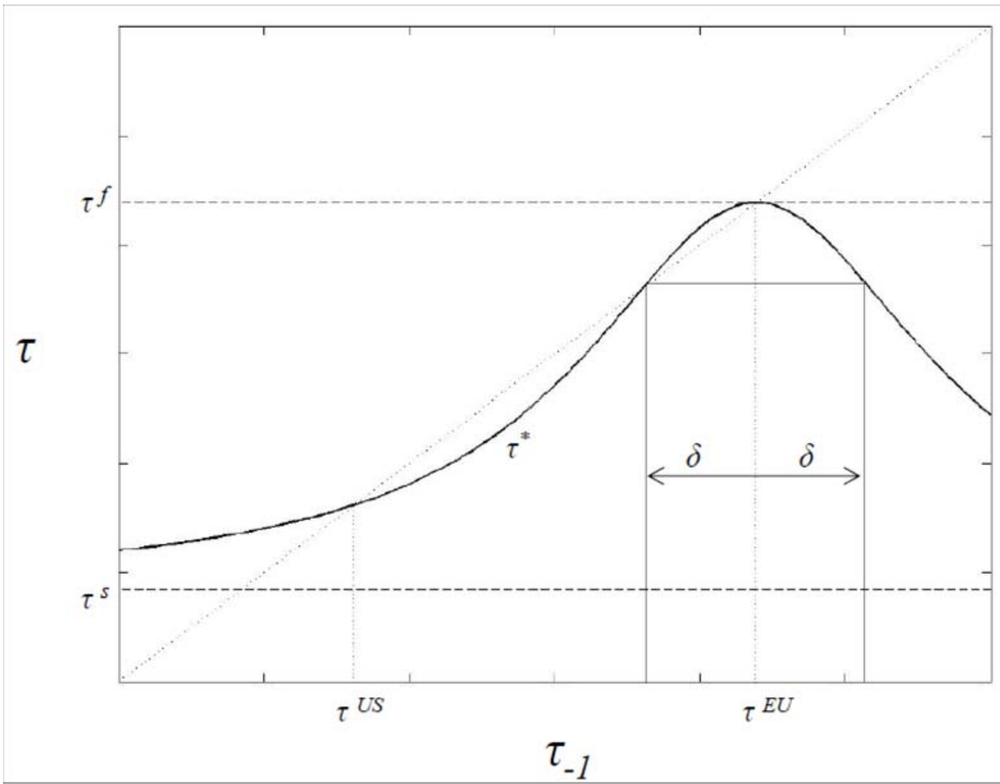

Figure 5: Multiple steady states when $\tau^f - \tau^s > 2\sqrt{\frac{\Phi}{2\bar{a}-a_{med}}}$

redistribution level lower than the fair level but higher than the selfish one, $\tau^s < \tau^* < \tau^f$.

According to our analysis, the high redistribution European style welfare state is then characterized by an emotional contagion whereas the low redistribution American style welfare system is characterized by a rational control: $\tau^s < \tau^{US} < \tau^{EU} = \tau^f$. As a consequence, in the European style system, as $\tau^{EU} = \tau^f$, any increase of the difference between mean and median incomes Δ can be associated with a decrease of the relative importance of luck in the wage determination L which tends to reduce the level of redistribution: $\frac{d\tau^{EU}}{d\Delta} < 0$. By contrast, we can show in the American style system that an increase of Δ is still associated with less redistribution, $\frac{d\tau^{US}}{d\Delta} > 0$, as in the Meltzer-Richard model (1981). This cultural analysis suggests then that we should dissociate Europe and the United States when testing the empirical relationship between redistribution and income inequality.

4 Family transfers and inequality

So far, in the income determination we have apprehended luck only as a white noise. By making this, we have obviously omitted one important dimension of luck which is being born in a wealthy and well-educated family, and therefore the importance of the bequests and any other inheritted social capital (status, network, ...) in the income determination. Let us now consider that individual life income can be characterized by:

$$y_{it} = \pi_{it-1} + a_{it} + e_{it} + \varepsilon_{it} \tag{17}$$

where π_{it-1} represents the parental transfers, bequests and social capital. If parental transfers are the expression of a *warm glow* effect or a *joy of giving*, i.e a selfish incentive, the transfers π_{it} that an adult will receive in $t+1$ enter in the utility of his parents in t, and we can redefine the rational-selfish utility as[8]:

$$U_{it} = \frac{1}{\psi^{\psi}(1-\psi)^{1-\psi}} c_{it}^{1-\psi} \pi_{it}^{\psi} - \frac{e_{it}^2}{2a_{it}} \tag{18}$$

It follows that maximizing (18) according to consumption and family transfers entails:

$$\begin{aligned} c_{it} &= (1-\psi)\left[(1-\tau_t) y_{it} + \tau_t \bar{y}_t\right] & (19) \\ \pi_{it} &= \psi\left[(1-\tau_t) y_{it} + \tau_t \bar{y}_t\right] & (20) \end{aligned}$$

and that introducing these optimal behaviors (19) and (20) into (18) leads to an unchanged rational utility compared to previous sections:

$$U_{it} = y_{it}(1-\tau_t) + \tau_t \bar{y}_t - \frac{e_{it}^2}{2a_{it}} \tag{21}$$

In this configuration, the level of effort which corresponds to the maximization of (21) is then still:

$$e_{it} = a_{it}(1-\tau_t) \tag{22}$$

[8]Even if we had consider a true altruistic motive, *loving his children* can be associated with nepotism which is a dynastic selfishness, and not with fairness or morality as stated by Sober and Wilson (1998).

Consider now that we can integrate in the concept of fair income the principle of *equal opportunities*. In this case we can redefine the fair level of income as:

$$\hat{y}_{it} = \bar{\pi}_{t-1} + a_{it} + e_{it}$$

This characterization of the fair level of income strongly differs from the one of Alesina and Angeletos (2005). The latter consider fairness under a dynastic basis, which means that they assume that a transfer received by a person is fair if related to an effort from his parents. Nevertheless, studies tend to support fairness on a more individual basis, i.e. that individuals tend to associate family transfers with unfairness. Alesina and La Ferrara (2005) show in particular that people who believe that family background influence income and then that opportunities in life are biaised favor redistribution in order to correct for unfair advantages.

If we assume that parental transfers π_{it-1} received by an adult in t are independent from his willpower and personal skills a_{it} and from his luck ε_{it}, and that distributions of a and ε are stable over time, the moral intuition which results from the minimization of (6) stays:

$$\tau_t^f = \begin{cases} 1 - \sqrt{1 - \frac{L_t}{2}} & \text{if } L_{t-1} \leq 2 \\ 1 & \text{otherwise} \end{cases}$$

where $L_t = \frac{\sigma^2_{\pi_{t-1}} + \sigma^2_\varepsilon}{\sigma^2_a}$, $\sigma^2_{\pi_{t-1}}$ being the variance of the family transfers received in t. Determination of his preferred tax rate by each individual corresponds then to the maximization in 1 of the utility minus the feeling of guilt, $E_1\left[y_{it}\left(1-\tau_t\right)+\tau\bar{y}_t-\frac{e^2_{it}}{2\beta_i}\right]-\frac{\Phi}{2\left[\tau^f_{t-1}-\tau^*_{t-1}\right]^2}\left(\tau^f_t-\tau_t\right)^2$, which results in:

$$\tau^*_t = \xi_{t-1}\tau^s_t + \left(1-\xi_{t-1}\right)\tau^f_t \tag{23}$$

where $\tau^s_t = \frac{2(\bar{a}-a_{med})+(\bar{\pi}_{t-1}-\pi_{medt-1})}{2\bar{a}-a_{med}}$ is the tax rate chosen by a median voter driven only by his self-interest, and $\xi_{t-1} = \frac{2\bar{a}-a_{med}}{2\bar{a}-a_{med}+\frac{\Phi}{\left[\tau^f_{t-1}-\tau^*_{t-1}\right]^2}}$.

Let us assume an institutional stationary history such that $\tau_s = \tau_{-1}, \forall s \leq t-1$, in such a case we have (see Appendix E):

$$\tau^s_t = \frac{\left[2+\frac{\psi(1-\tau_{-1})(2-\tau_{-1})}{1-\psi(1-\tau_{-1})}\right]\left(\bar{a}-a_{med}\right)}{2\bar{a}-a_{med}}$$

$$\sigma^2_\pi\left(\tau_{-1}\right) = \frac{\psi^2\left(1-\tau_{-1}\right)^2}{1-\psi^2\left(1-\tau_{-1}\right)^2}\left[\left(2-\tau_{-1}\right)^2\sigma^2_a+\sigma^2_\varepsilon\right]$$

Proposition 5 *Under a stationnary tax history $\{\tau_s = \tau_{-1}\}^{s=t-1}_{s=-\infty}$, any stationary level of taxation only driven by moral intuitions $\tau_t = \tau^f$ or by self-interest $\tau_t = \tau^s$ is unique.*

Proposition 6 *Under a stationnary tax history $\{\tau_s = \tau_{-1}\}^{s=t-1}_{s=-\infty}$, if $\tau^f - \tau^s$ is sufficiently high, the model exhibits two steady states such that $\tau^s < \tau^{US} < \tau^{EU} = \tau^f$.*

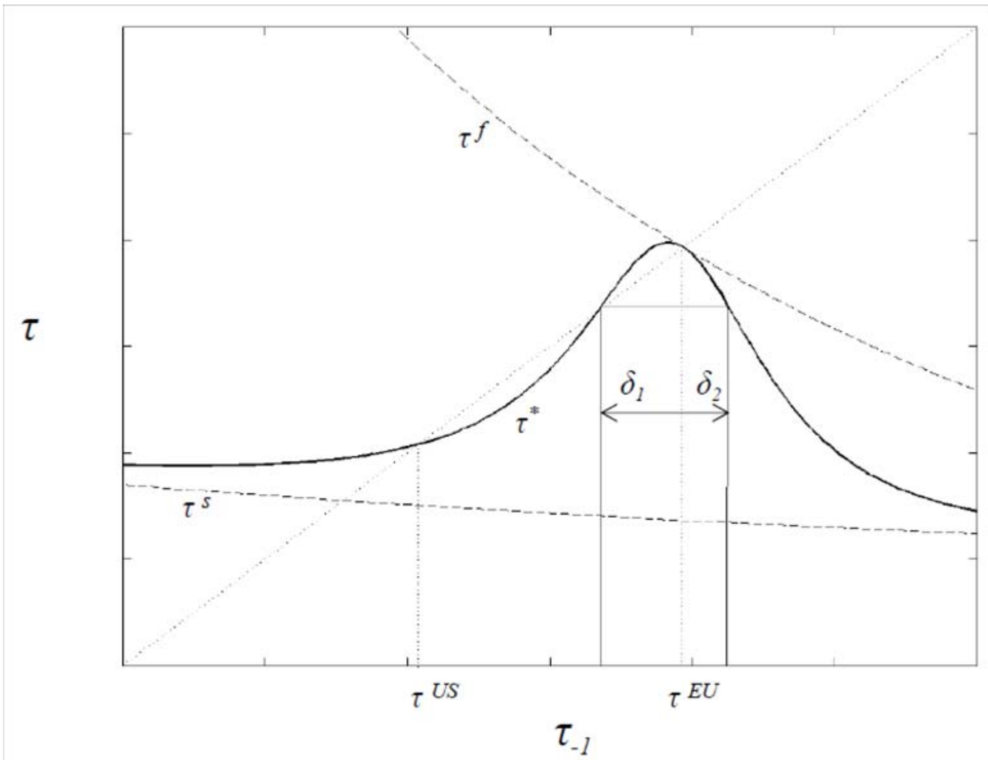

Figure 6: Multiple steady states and family history

From proposition 5 et 6, again we can assert that the possibility of a multiple steady states results from the interaction of the two distinct processes. We can not detail explicitly the conditions of existence as in the previous section, but by similarity it is easy to deduce that a multiplicity can arise only if τ^f and τ^s are sufficiently differents. A limit case where it is proven is of course when ψ is close to 0. Indeed, in such a case we are without family transfers, i.e. in the case treated in the previous section. Considering family transfers allow still us to interpret the difference of redistribution between Europe and the United States as a distinction in the nature of the leading cognitive process when voting: emotional (system 1) in Europe, rational (system 2) in the United States.

5 Conclusion

If we consider that humans are only driven by their self interest, Meltzer and Richard (1981) show that the level of redistribution in a democratic society is growing with the inequality of the income distribution. A result which is weakly supported by the data. In this article, we assert that this failure of the canonical model can be in particular associated with its behavioral assumption. Modern cognitive sciences along with empirical studies converge in one major point: morality and altruism are essentials to explain redistribution. In the line of a "new synthesis" in moral psychology, we then assert that the voting behavior over redistribution is best characterized by first an automatic cognitive process which generates quicly intuitions on the fair level of redistribution, and second by a rational self oriented reasoning which controls the feeling of guilt associated with the fair intuitions. As a result of this dual-process decision-making, we show that the U-shape between inequality and redistribution supported by the data is a general feature of the model. In addition, assuming that the feeling of guilt is context dependent and is reduced if the previous generation failed in implementing the intuitively fair institution, the model exhibits a multiplicity of steady states which can explain the huge difference of redistribution observed between Europe and the United States.

The approach we use in this article raises two issues which can lead to further research. First, the mind architecture that we use in this article, in the line of Kahneman (2003) and Haidt (2001), states a clear partition between an emotional process and a rational one. For most neurobiologists, such a partition is a highly stylised distinction with no strong basis. By contrast, in the neural workspace model proposed by Dehaene et al. (1998) and Dehaene and Naccache (2001), lots of distinct groups of neurons convey in parrallel different representation of the external world, and the resulting conscious perception of the external world often adopts the information of one neural group and entirely suppresses the information carried by the others according to the *winner-take-all* principle (Camerer et al., 2005). Second, our approach suggests that emotions are by essence morals. However, as well as guilt we can think of greed as an important emotion shaping human behavior, especially when considering voting over redistribution.

References

[1] Alesina A., Glaeser E. and Sacerdote B. (2001), Why doesn't the US have a European-style welfare system?, *Brookings Papers on Economic Activity*, 2, pp. 187-277.

[2] Alesina A. and Glaeser E. (2004), *Fighting poverty in the US and Europe: A world of difference*, Oxford University Press.

[3] Alesina A. and Angeletos G.-M. (2005), Fairness and redistribution: US versus Europe, *American Economic Review,* 95(4), pp. 960-980.

[4] Alesina A. and La Ferrara E. (2005), Preferences for redistribution in the land of opportunities, *Journal of Public Economics*, 89(5-6), pp. 897-931.

[5] Alesina A. and Giuliano P. (2009), Preferences for redistribution, *IZA DP* no 4056, March.

[6] Anderson S., Bechara A., Damasio H., Tranel D. and Damasio A. (1999), Impairment of social and moral behavior related to early damage in human prefrontal cortex, *Nature Neuroscience*, 2, pp. 1032-1037.

[7] Batson C. D. (1991), *The altruism question: toward a social psychological answer*, Hillsdale, NJ: Lawrence Erlbaum.

[8] Bechara A. (2004), The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage, *Brain and Cognition*, 55, pp. 30-40.

[9] Bénabou R. (2000), Unequal societies: income distribution and the social contract, *American Economic Review*, 90, pp. 96-129.

[10] Bénabou R. and Ok E. (2001), Social mobility and the demand for redistribution: the POUM hypothesis, *Quarterly Journal of Economics*, 116(2), pp. 447-487.

[11] Bénabou R. and Tirole J. (2003), Intrinsic and extrinsic motivations, *Review of Economic Studies*, 70, pp. 489-520.

[12] Bénabou R. and Tirole J. (2006), Belief in a just world and redistributive politics, *Quarterly Journal of Economics*, 121(2), pp. 699-746.

[13] Berthoz S., Armony J., Blair R. and Dolan R. (2002), An fMRI study of intentional and unintentional (embarrassing) violations of social norms, *Brain*, 125, pp. 1696–708.

[14] Berthoz S., Grezes J., Armony J., Passingham R. and Dolan R. (2006), Affective response to one's own moral violations, *NeuroImage*, 31(2), pp. 945-950.

[15] Bjorklund A. and Jantti M. (1997), Intergenerational income mobility in Sweden compared to the United States, *American Economic Review*, 87(5), pp. 1009-1018.

[16] Blair R. (2001), Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy, *Journal of Neurology, Neurosurgery & Psychiatry*, 71(16), pp. 727-731.

[17] Campante F. (2006), Redistribution in a model of voting and campaign contributions, mimeo, Harvard University.

[18] Camerer C., Loewenstein G. and Prelec D. (2005), Neuroeconomics: how neuroscience can informe economics, *Journal of Economic Literature*, p. 9-64.

[19] Cohen J. (2005), The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion, *Journal of Economic Perspectives*, 19(4), pp. 3-24.

[20] Corneo G. and Grüner H.-P.(2002), Individual preferences for political redistribution, *Journal of Public Economics*, 83, pp. 83-107.

[21] Corneo G. and Fong C. (2008), What's the monetary value of distributive justice?, *Journal of Public Economics*, 92(1), pp. 289-308.

[22] Damasio A. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, Penguin Books.

[23] Decety J. and Chaminade T. (2003), Neural correlates of feeling sympathy, *Neuropsychologia*, 41, pp. 127-138.

[24] Dehaene S., Kerszberg M. and Changeux J.P. (1998), A neuronal model of a global workspace in effortful cognitive tasks, *Proceedings of the National Academy of Sciences USA*, 95, pp. 14529-14534.

[25] Dehaene S. and Naccache L. (2001), Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework, *Cognition*, 79, pp. 1-37.

[26] de Mello L. et Tiongson E. (2006), Income inequality and redistributive government spending, *Public Finance Review*, 34(3), pp. 282-305.

[27] Edelstein R. and Shaver P. (2007), A cross-cultural examination of lexical studies of self-conscious emotions, in Tracy J., Robins R. and Tangney J. P. (Eds), *The self-conscious emotions: Theory and research*, NY: The Guilford Press, chap. 11.

[28] Eisenberg N. (2000), Emotion, regulation and moral development, *Annual Review of Psychology*, 51, pp. 665-697.

[29] Ekman P. and Friesen W. V. (1971), Constants across cultures in the face and emotion, *Journal of Personality and Social Psychology*, 17, pp. 124-129.

[30] Ekman P. (1992), An argument for basic emotions, *Cognition and Emotion*, 6, pp. 169-200.

[31] Evans J. (2008), Dual-processing accounts of reasoning, judgment and social cognition, *Annual Review of Psychology*, 59, pp. 255-278.

[32] Fehr E. and Schmidt K. (2006), The economics of fairness, reciprocity and altruism: experimental evidence and New Theories, in S.-C. Kolm and J. Mercier Ythier (Eds), *Handbook of the economics of giving, altruism and reciprocity, vol. 1*, North Holland, chap. 8.

[33] Fong C. (2001), Social preferences, self-interest, and the demand for redistribution, *Journal of Public Economics*, 82(2), pp. 225-246.

[34] Forsé M. and Parodi M. (2006), Justice distributive: la hiérachie des principes selon les européens, *Revue de l'OFCE*, 98, pp. 213-244.

[35] Frijda N. and Mesquita B. (1994), The social roles and functions of emotions, in Kitayama S. and Markus H. (Eds), *Emotion and Culture: Empirical studies of mutual influence*, Washington DC: American Psychological Association, chap. 3.

[36] Gilens M. (1999), *Why Americans hate welfare*, University of Chicago Press, Chicago.

[37] Goetz J. and Keltner D. (2007), Shifting meanings of self-conscious emotions across cultures: A social-functional approach, in Tracy J., Robins R. and Tangney J. P. (Eds), *The self-conscious emotions: Theory and research*, NY: The Guilford Press, chap. 9.

[38] Gray J. (2004), Integration of emotion and cognitive control, *Current Directions in Psychological Science*, 13, pp. 46-48.

[39] Greene J., Sommerville R., Nystrom L., Darley J., Cohen J. (2001), An fMRI investigation of emotional engagement in moral judgement, *Science*, 293, pp. 2105-2108.

[40] Greene J., Nystrom L., Engell A., Darley J., Cohen J. (2004), The moral bases of cognitive conflict and control in moral judgment, *Neuron*, 44, pp. 389-400.

[41] Greene, J. (2005), Cognitive neuroscience and the structure of the moral mind, in S. Laurence, P. Carruthers. and S. Stich. (Eds.), *The Innate Mind: Structure and Contents*, Oxford University Press, New York.

[42] Güth W., Schmittberger R. and Schwarze B. (1982), An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior & Organization*, 3, p. 367-388.

[43] Haidt J. (2001), The emotional dog and its rational tail: a social intuitionist approach to moral judgment, *Psychological Review*, 108, pp. 814-834.

[44] Haidt J. (2007), The new synthesis in moral psychology, *Science*, 316(5827), p.998-1002.

[45] Haidt J. (2008), Morality, *Perspectives on Psychological Science*, 3, p.65-72.

[46] Harlé M. and Sanfey A. (2007), Incidental sadness biases social economic decisions in the Ultimatum Game, *Emotion*, 7(4), pp. 876-881.

[47] Hauser M. (2006), *Moral minds: How nature design our universal sense of right and wrong*, Harper Collins.

[48] Henrich J. (2000), Does culture matter in economic behavior? Ultimatum Game bargaining among the Machiguenga of the Peruvian Amazone, *American Economic Review*, 90(4), pp. 973-979.

[49] Henrich J., Boyd R., Bowles S., Camerer C., Gintis H., McElreath R. and Fehr E. (2001), In search of Homo economicus: Experiments in 15 small-scale societies, *American Economic Review*, 91(2), pp. 73-79.

[50] Henrich J., Boyd R., Bowles S., Gintis H., Fehr E., Camerer C., McElreath R., Gurven M., Hill K., Barr A., Ensminger J., Tracer D., Marlow F., Patton J., Alvard M., Gil-White F. and Henrich N. (2005), "Economic Man" in cross-cultural perspective: Ethnography and experiments from 15 small-scale societies, *Behavioral and Brain Studies*, 28, pp. 795-855.

[51] Hsu M., Anen C. and Quartz S. (2008), The right and the good: Distributive justice and neural encoding of equity and efficiency, *Science*, 320, pp. 1092- 1095.

[52] Iversen T. and Soskice D. (2006), Electoral institutions and the politics of coalitions: why some democracies redistribute more than others, *American Political Science Review*, 100(2), pp. 165-181.

[53] Kahneman D. (2003), Maps of bounded rationality: Psychology for behavioral economics, *American Economic review*, 93(5), pp. 1449-1475.

[54] Koenigs M. and Tranel D. (2007), Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum

Game, *The Journal of Neuroscience*, 27(4), pp. 951-956.

[55] Kohlberg L. (1981), *The philosophy of moral development: Moral stages and the idea of justice*, vol 1, Harpers & Row, San Francisco.

[56] LeDoux J. (1996), *The Emotional Brain*, New York, Simon and Schuster.

[57] Lerner M. (1982), *The Belief in a Just World: a fundamental delusion*, New York, NY: Plenum Press.

[58] Lind J. T. (2005), Why is there so little redistribution?, *Nordic Journal of Political Economy*, 31, pp. 111-125.

[59] Lind J. T. (2007),

[60] Markus H. and Kitayama S. (1991), Culture and the self: implications for cognition, emotion, and motivation, *Psychological Review*, 98(2), pp. 224-253

[61] Meltzer A. and Richard S. (1981), A rational theory of the size of government, *Journal of Political Economy*, 89(5), pp 914-927.

[62] Mesquita B. and Frijda N. (1992), Cultural variations in emotions: A review, *Psychological Bulletin*, 112(2), pp. 179-204.

[63] Mikhail J. (2007), Universal moral grammar: theory, evidence and the future, *Trends in Cognitive Sciences*, 11(4), pp. 143-152.

[64] Moene K. O. et Wallerstein (2001), Inequality, social insurance and redistribution, *American Political Science Review*, 95(4), pp. 859-874.

[65] Moll J., Andreiuolo P.A., de Oliveira-Souza R., Eslinger P., Mourão-Miranda J. and Pessoa L. (2002), The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions, *The Journal of Neuroscience*, 22(7), pp. 2730-2736.

[66] Moll J., Zahn R., de Oliveira-Souza R., Krueger F.and Grafman J. (2005), The neural basis of human moral cognition, *Nature Reviews Neuroscience*, 6, pp. 799-809.

[67] Nado J., Kelly D and Stich S. (2006), Moral judgement, in Symons J. and Calvo P. (Eds), *Routledge Companion to the Philosophy of Psychology*, Routledge, forthcoming.

[68] Perotti R. (1996), Growth, income distribution and democracy: what the data say, *Journal of Economic Growth*, 1(2), pp. 149–187.

[69] Pessoa L. (2008), On the relationship between emotion and cognition, *Nature Reviews Neuroscience*, 9, pp. 223-227.

[70] Petrova M. (2008), Inequality and Media Capture, *Journal of Public Economics*, 92(1-2), pp. 183-212.

[71] Phelps E. (2009), The study of emotion in neuroeconomics, in Glimcher P., Camerer C., Fehr E. and Poldrack R. (Eds), *Neuroeconomics: Decision making and the brain*, Academic Press, Elsevier, chap. 16.

[72] Piketty T. (1995), Social mobility and redistributive politics, *Quarterly Journal of Economics*, 110(3), pp. 551-584.

[73] Piketty T. (2003), Attitudes vis-à-vis des inégalités en France : existerait-il un consensus ?, *Comprendre*, no 4, p. 209-241.

[74] Rodriguez F. (2004), Inequality, redistribution and rent-seeking, *Economics & Politics*, 16, pp. 287-320.

[75] Ryan R. and Deci E. (2000), Intrinsic and extrinsic motivations: classic definitions and new directions, *Contemporary Educational Psychology*, 25, pp. 54-67.

[76] Sanfey A., Rilling J., Aronson J., Nystrom L. and Cohen J. (2003), The neural basis of economic decision-making in the ultimatum game, *Science*, 300, pp. 1755-1758.

[77] Scherer K. (1997), The role of culture in emotion-antecedent appraisal, *Journal of Personality and Social Psychology*, 73(5), pp. 902-922.

[78] Schokkaert E. (1998), Mr. Fairmind is post-welfarist: opinions on distributive justice, *Center for Economic Studies DPS 98.09*, Catholic University of Leuven.

[79] Sober E. and Wilson D. S. (1998), *Unto others: The evolution and psychology of unselfish behavior*, Harvard University Press, Cambridge, MA.

[80] Roemer J. (1998), Why do the poor do not expropriate the rich: an old argument in new garb, *Journal of Public Economics*, 70, pp. 399-424.

[81] Tracy J. and Robins R. (2004), Putting the self into self-conscious emotions: a theoretical model, *Psychological Inquiry*, 15(2), pp. 103-125.

[82] Twenge J., Baumeister R., DeWall N., Ciarocco N. and Bartels M. (2007), Social exclusion decreases prosocial behavior, *Journal of Personality and Social Psychology*, 92(1), pp. 56-66.

[83] van 't Wout M., Kahn R., Sanfey A. and Aleman A. (2006), Affective state and decision-making in the Ultimatum Game, *Experimental Brain Research*, 169, pp. 564-568.

[84] Young L. and Koenigs M. (2007), Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology, *British Medical Bulletin*, 84, pp. 69-79.

Appendix A. Clusters of attributes associated with dual-process decision-making (from Evans, 2008)

System 1	System 2
Cluster 1 (Consciousness)	
Unconscious (preconscious)	Conscious
Implicit	Explicit
Automatic	Controlled
Low effort	High effort
Rapid	Slow
High capacity	Low capacity
Default process	Inhibitory
Holistic, perceptual	Analytic, reflective
Cluster 2 (Evolution)	
Evolutionarily old	Evolutionarily recent
Evolutionary rationality	Individual rationality
Shared with animals	Uniquely human
Nonverbal	Linked to language
Modular cognition	Fluid intelligence
Cluster 3 (Functional characteristics)	
Associative	Rule based
Domain specific	Domain general
Contextualized	Abstract
Pragmatic	Logical
Parallel	Sequential
Stereotypical	Egalitarian
Cluster 4 (Individual differences)	
Universal	Heritable
Independent of general intelligence	Linked to general intelligence
Independent of working memory	Limited by working memory capacity

Appendix B. The fair-oriented cognitive process

The moral objective is characterized by

$$F=\int_i \left\{\left[(1-\tau)\,y_i+\tau\bar{y}\right]-\hat{y}_i\right\}^2 di$$

According to eq. (4),

$$F=\int_i \left\{\left[(1-\tau)\left(a_i\,(2-\tau)+\varepsilon_i\right)+\tau\bar{a}\,(2-\tau)\right]-a_i\,(2-\tau)\right\}^2 di$$
$$=\int_i \left\{(1-\tau)\,\varepsilon_i-\tau\,(2-\tau)\,(a_i-\bar{a})\right\}^2 di.$$

As α and ε are independently distributed over the population,

$$F=(1-\tau)^2\int_i \varepsilon_i^2 di+\tau^2\,(2-\tau)^2\int_i (a_i-\bar{a})^2\,di=(1-\tau)^2\,\sigma_\varepsilon^2+\tau^2\,(2-\tau)^2\,\sigma_a^2.$$

It follows that

$$\frac{F}{\sigma_a^2}=(1-\tau)^2\,L+\tau^2\,(2-\tau)^2$$

where $L=\frac{\sigma_\eta^2}{\sigma_\alpha^2}$.

Appendix C. Proof of proposition 2

$$\frac{\partial\tau^*}{\partial(\bar{\alpha}-\alpha_{med})}=\frac{\partial\tau^*}{\partial\sigma_\alpha^2}\left[\frac{\partial(\bar{\alpha}-\alpha_{med})}{\partial\sigma_\alpha^2}\right]^{-1}\Rightarrow sign\left[\frac{\partial\tau^*}{\partial(\bar{\alpha}-\alpha_{med})}\right]=sign\left[\frac{\partial\tau^*}{\partial\sigma_\alpha^2}\right]$$

Differentiating eq. (13) gives:

$$\frac{\partial\tau^*}{\partial\sigma_\alpha^2}=\frac{\partial\xi}{\partial\sigma_\alpha^2}\tau^s+\xi\frac{\partial\tau^s}{\partial\sigma_\alpha^2}-\frac{\partial\xi}{\partial\sigma_\alpha^2}\left[1-\left(1-\frac{1}{2}\frac{\sigma_\eta^2}{\sigma_\alpha^2}\right)^{\frac{1}{2}}\right]+[1-\xi]\left[\frac{-\sigma_\eta^2}{4\left(1-\frac{1}{2}\frac{\sigma_\eta^2}{\sigma_\alpha^2}\right)^{\frac{1}{2}}\sigma_\alpha^4}\right],$$

where $\frac{\partial\xi}{\partial\sigma_\alpha^2}=\frac{\left(2\frac{\partial\bar{\alpha}}{\partial\sigma_\alpha^2}-\frac{\partial\alpha_{med}}{\partial\sigma_\alpha^2}\right)\varphi}{\left(2\bar{\alpha}-\alpha_{med}+\varphi\right)^2}>0.$

It follows that:

$$\lim_{\sigma_\alpha^2\to\infty}\frac{\partial\tau^*}{\partial\sigma_\alpha^2}=\frac{\partial\xi}{\partial\sigma_\alpha^2}\tau^s+\xi\frac{\partial\tau^s}{\partial\sigma_\alpha^2}>0$$

and $\lim_{\sigma_\alpha^2 \to \frac{\sigma_\eta^2}{2}} \frac{\partial \tau^*}{\partial \sigma_\alpha^2} = -\infty$.

Appendix D. Proof of proposition 3

Let us define $\delta_t = \tau^f - \tau_t^*$ the difference between the fair and the effective level of taxation, eq (15) can be rewritten as:

$$\delta_t = \frac{2\bar{a} - a_{med}}{2\bar{a} - a_{med} + \frac{\Phi}{\delta_{t-1}^2}} \left(\tau^f - \tau^s\right) \tag{24}$$

and the stationnarity is then defined by:

$$\delta^3 - \left(\tau^f - \tau^s\right)\delta^2 + \frac{\Phi}{2\bar{a} - a_{med}}\delta = 0 \tag{25}$$

If $\left|\tau^f - \tau^s\right| > 2\sqrt{\frac{\Phi}{2\bar{a}-a_{med}}}$, eq. (25) exhibits three roots $\delta = 0$, $\delta = \frac{\tau^f-\tau^s+\sqrt{\left(\tau^f-\tau^s\right)^2-\frac{4\Phi}{2\bar{a}-a_{med}}}}{2}$ and $\delta = \frac{\tau^f-\tau^s-\sqrt{\left(\tau^f-\tau^s\right)^2-\frac{4\Phi}{2\bar{a}-a_{med}}}}{2}$.

In addition, as $\lim_{\delta^2 \to 0} \frac{\partial\left[\frac{2\bar{a}-a_{med}}{2\bar{a}-a_{med}+\frac{\Phi}{\delta^2}}\right]}{\partial \delta^2} = 0$, the dynamics can be represented as in Figure (7) if $\tau^f > \tau^s$. Therefore, in that case there exists two stable equilibria characterized by $\delta = 0$ and $\delta = \frac{\tau^f-\tau^s+\sqrt{\left(\tau^f-\tau^s\right)^2-\frac{4\Phi}{2\bar{a}-a_{med}}}}{2}$, i.e. respectively by $\tau^* = \tau^f$ and $\tau^* = \frac{1}{2}\left(\tau^f + \tau^s - \sqrt{\left(\tau^f - \tau^s\right)^2 - \frac{4\Phi}{2\bar{a}-a_{med}}}\right)$.

Appendix E. Stationary history

From (17), (20) and (22) we have:

$$\pi_{it-1} = \psi\left\{(1-\tau_{t-1})\left[\pi_{it-2} + (2-\tau_{t-1})\,a_{it-1} + \varepsilon_{it-1}\right] + \tau_{t-1}\left[\bar{\pi}_{t-2} + (2-\tau_{t-1})\,\bar{a}_{t-1}\right]\right\} \tag{26}$$

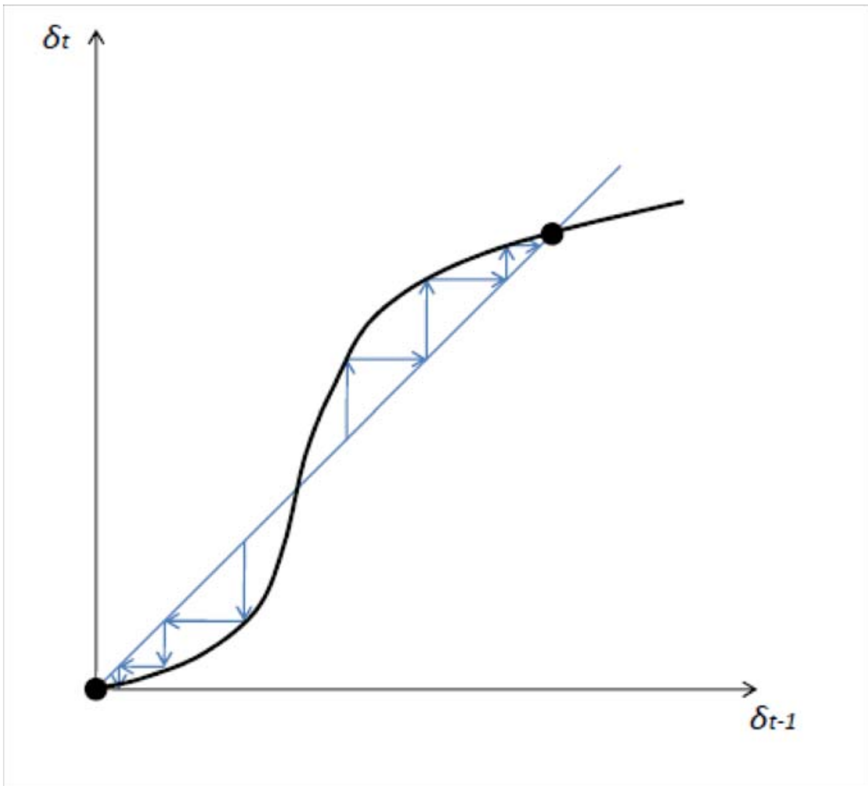

Figure 7: Dynamics and stable equilibria when $\tau^f - \tau^s > 2\sqrt{\frac{\Phi}{2\bar{a} - a_{med}}}$

Let us assume an institutional stationary history such that $\tau_s = \tau_{-1}, \forall s \leq t-1$, in such a case, from (26) we have $\sigma^2_{\pi_{t-1}} = \sigma^2_\pi(\tau_{-1})$

$= \psi^2 (1-\tau_{-1})^2 \left[\sigma^2_\pi(\tau_{-1}) + (2-\tau_{t-1})^2 \sigma^2_a + \sigma^2_\varepsilon\right]$ and therefore:

$$\sigma^2_\pi(\tau_{-1}) = \frac{\psi^2 (1-\tau_{-1})^2}{1-\psi^2 (1-\tau_{-1})^2} \left[(2-\tau_{-1})^2 \sigma^2_a + \sigma^2_\varepsilon\right]$$

It follows that $\frac{\partial \sigma^2_\pi(\tau_{-1})}{\partial \tau_{-1}} < 0$ and then that $\frac{\partial \tau^f_t}{\partial \tau_{-1}} \leq 0$.

From (26) we also have

$\bar{\pi}_{t-1} = \bar{\pi}(\tau_{-1}) = \frac{\psi}{1-\psi(1-\tau_{-1})} \{(1-\tau_{-1})(2-\tau_{-1})\bar{a} + \tau_{-1}[\bar{\pi}(\tau_{-1}) + (2-\tau_{-1})\bar{a}]\}$

and

$\pi_{medt-1} = \pi_{med}(\tau_{-1}) = \frac{\psi}{1-\psi(1-\tau_{-1})} \{(1-\tau_{-1})(2-\tau_{-1}) a_{med} + \tau_{-1}[\bar{\pi}(\tau_{-1}) + (2-\tau_{-1})\bar{a}]\}$.

It follows that:

$$\bar{\pi}_{t-1} - \pi_{medt-1} = \bar{\pi}(\tau_{-1}) - \pi_{med}(\tau_{-1}) = \frac{\psi(1-\tau_{-1})(2-\tau_{-1})}{1-\psi(1-\tau_{-1})}(\bar{a} - a_{med})$$

and therefore:

$$\tau^s_t = \frac{2(\bar{a} - a_{med}) + (\bar{\pi}_{t-1} - \pi_{medt-1})}{2\bar{a} - a_{med}} = \frac{\left[2 + \frac{\psi(1-\tau_{-1})(2-\tau_{-1})}{1-\psi(1-\tau_{-1})}\right](\bar{a} - a_{med})}{2\bar{a} - a_{med}}$$

It follows that $\frac{\partial \tau^s_t}{\partial \tau_{-1}} \leq 0$.