

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

<http://www.archive.org/details/socialmobilityre00pike>

HB31
.M415

no.
94-
15

56

Dewey

**working paper
department
of economics**

Social Mobility and Redistributive Politics

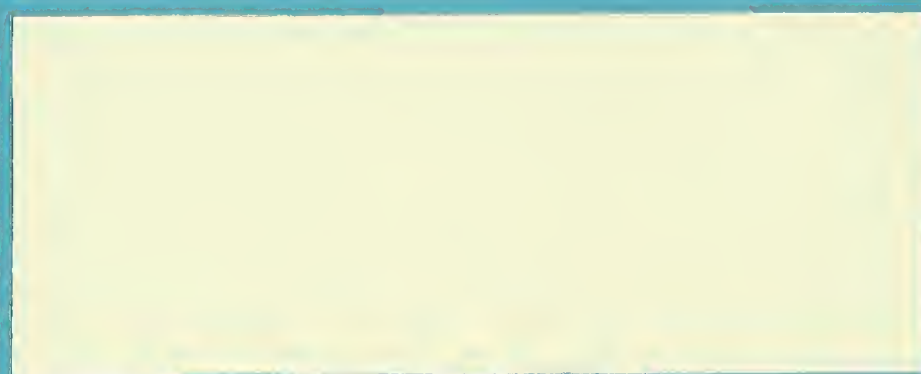
Thomas Piketty

94-15

April 1994

**massachusetts
institute of
technology**

**50 memorial drive
cambridge, mass.02139**



Social Mobility and Redistributive Politics

Thomas Piketty

94-15

April 1994



M.I.T. LIBRARIES

JUL 12 1994

RECEIVED

january 1994 (revised april 1994).

Social Mobility and Redistributive Politics*

**Thomas Piketty
Economics Dept. - MIT
Cambridge, MA 02139, USA**

Abstract : Assume agents differ in their beliefs concerning the relative importance of individual effort and social rigidities in shaping individual achievements, so that they trade off differently the social benefits of equalizing opportunities with the incentive costs of taxation (just as economists). Through their individual mobility experience they are exposed to different signals regarding these structural parameters, and in the long-run "left-wing dynasties" believing less in individual effort and voting for more redistribution coexist with "right-wing dynasties". Thus we are able to explain (1) why rich and poor claim the same abstract principles of distributive justice but vote differently, (2) why social origins and not only current income play a crucial but (mostly) indirect role in shaping one's political attitudes, (3) how persistent differences in popular beliefs about social mobility and the need for redistribution can sustain although the underlying, "true" mobility rates are essentially the same (US vs Europe), (4) why individual countries tend to be politically homogeneous.

***I am grateful to seminar participants at MIT, Harvard (Economics Dept. and Kennedy School), Columbia and Boston University for their comments.**

Section 1 : Introduction.

This paper develops a rational-choice theory¹ of redistributive politics seeking to explain important stylised facts concerning the effect of social mobility both on individual political attitudes and aggregate political outcomes.

The idea that social mobility plays a crucial role in shaping political attitudes (in particular towards redistribution) has a long history in the social sciences. Tocqueville(1835) first stressed the idea that the difference in attitudes toward redistribution between Europe and the United States could be explained by presumed differences in mobility rates. Since then, many authors have followed this line to explain the absence of any strong socialist movement in the US, among which Marx(1852), Sombart(1906) and Petersen(1953). On the other hand comparative empirical studies of social mobility rates have long demonstrated the absence of any significant difference between industrial nations (see, e.g., Lipset-Bendix(1959), Erikson-Goldthorpe(1985, 1992)). Lipset-Bendix(1959) and Lipset(1966,1977,1992) have repeatedly suggested that persistent differences between European and US redistributive politics may be due to persistent differences in popular beliefs about social mobility².

¹That is, as we understand it, a theory describing precisely the values and preferences individuals are promoting, the information sets they are exposed to and the institutions aggregating their actions. This differs from most sociological "explanations" of the effect of one's mobility experience on one's political attitudes.

² "What explains the contrast in the political values and allegiances of American workers with those of other democratic nations? (...) the belief system concerning class rigidities stemming from varying historical experiences (...) seems much more important than slight variations in rates of mobility". [Lipset(1992, pp.xx-xxi)]. Regarding the presumed lesser importance of government as a redistributor of income in the US as

But social mobility is known to have crucial effects at the individual level as well. Although current income is positively correlated with voting attitudes toward redistribution (higher-income groups vote less for left-wing redistributive policies), the correlation is much less than one, and upwardly- or downwardly-mobile voters always exhibit an intermediate position between stable low-income and high-income voters³; that is, the following table summarizes the typical voting patterns observed across time and industrial democracies (see, e.g., Abramson(1973), Thomson(1971), Boy(1980), Cherkaoui(1992)):

		respondent's income ⁴	
		low-income	high-income
parents' income	low-income	70%	40%
	high-income	45%	25%

Table 1. Percentage of votes for left-wing parties as a function of individual mobility experience.

That is, seven out of ten lower-class voters born in the lower-class typically vote for left-wing parties, against less than one half of lower-class voters born in the middle-class. From this matrix it would appear that parents' income class determines one's political attitudes as

compared to Europe, see, e.g., the table reported in Mueller(1989, p.336).

³A few studies found that upwardly-mobile agents are on average more right-wing than stable middle-class (mostly in the US); however later studies have shown that this was non-robust (see Thomson(1971)) and this thesis has apparently been abandoned.

⁴This sociology/political science literature usually cuts the society into two halves: lower-class, manual occupations, and middle-class, non-manual occupations. Although this is highly rudimentary, more sophisticated studies confirm the basic findings (see Turner(1992)).

much as one's current income, whereas straight economic rationality should imply that only current income and not past family income⁵ should determine one's interests in redistribution⁶, as in the standard economic models of redistributive politics⁷.

Our primary objective is to provide a common framework to account for these various stylised facts. The basic idea of our theory is that although voters may share common distributive goals, through their various mobility experiences they (rationally) happen to learn and to believe different things concerning the openness of their society and how serious the incentive problem is. That is, we model rational agents as trying to learn from their income trajectory not only the mobility matrix of their society but mostly how responsive individual promotions and achievements are to individual effort (as opposed to predetermined factors), so as to evaluate the incentive costs of redistributive taxation. Such a learning process is essentially of the same nature as Rothschild(1974)'s multi-armed

⁵To the extent that the process of intergenerational income mobility exhibits no memory, which seems reasonable.

⁶One could argue that not only redistribution is involved when voting for some political party. However the picture survives when disaggregated studies try to isolate for attitudes toward inequality and redistribution. See the studies edited by Miller(1992).

⁷See, e.g., Mueller(1989) for the standard economic models of redistributive politics, and Perotti(1992). Aside from the stylised facts mentioned above (which by nature these theories cannot accommodate), the usual median-voter model of redistribution does not seem to be particularly consistent with the data (see, e.g., Alesina and Perotti(1993) and Perotti(1992)). See Piketty(1993) for an alternative viewpoint on the political economy of redistribution with perfectly-informed, selfish voters.

bandit problem⁸, and in the same way, costly experimentation implies that different dynasties converge toward different beliefs regarding society's mobility parameters and therefore different beliefs concerning the socially-optimal redistribution rate.

The key point is that in the long-run, the same reasons lead some dynasties to support higher taxation and redistribution and at the same time to supply less effort, while some other dynasties support lower redistribution and at the same time work harder to be successful; namely, in the long-run some dynasties believe (maybe rightly) that predetermined factors are more important than individual effort in shaping individual achievements, while some others believe (maybe rightly) that individual effort is the key to success and social rigidities are second-order⁹. This implies that in steady-state there are more "left-wing dynasties" in the lower-class and more "right-wing dynasties" in the middle-class (regardless of which dynasties have the "right" beliefs, if any), although everybody started with the same distributive goal. Moreover, upwardly- and downwardly-mobile groups include intermediate fractions of left-wing and right-wing dynasties as compared to stable lower-class and upper-class agents, which leads exactly to the voting patterns depicted in table 1.

The multiplicity of steady-states explains at the same time why different countries can remain in different redistributive equilibria although the underlying structural parameters

⁸The learning process under consideration is actually more sophisticated than a standard multi-armed bandit problem, both because individual learning depends on some aggregate variable (redistributive taxation) and because of the possibility of learning from others' experiences; however this does not change the basic result of long-run heterogeneity (see below, and especially section 4).

⁹In fact, there's a all continuum in between these two extreme dynasties.

of mobility are essentially the same. This is particularly likely if a country exhibited for some time in the past a significantly different experience of social mobility before joining the "common" pattern (the 19th century US had a significantly different social structure before converging to common features along with Europe).

Four different pieces of evidence lead us to think this theory has some relevance. First, when asked what they think about inequality and redistribution and why they vote the way they do, it appears that people from different social backgrounds share a wide consensus about abstract principles of distributive justice (ability per se is usually considered as an irrelevant basis for desert unless it is seen as being a result of previous efforts; people can deserve unequal rewards only on the basis of features (such as effort) that are subject to voluntary control), but that they differ substantially on practical assessments concerning the key to personal success (the poor emphasizing structural factors, the rich, personal qualities such as effort and ambition) (see Rytina, Form and Pease(1970), Kluegel and Smith(1986, chaps.3-4), and Miller(1992)). In some sense, this paper chooses to take seriously people's justification of their attitudes toward redistribution, instead of describing them as egoistic and liar from the beginning¹⁰.

Next, voting patterns exhibit indeed an amazingly high rate of dynastic reproduction: Abramson(1972) shows Italian data where more than 80% of voters with left-wing parents

¹⁰One could obviously argue that people are basically egoistic and ex-post "find" some beliefs to justify their behavior. But then one has to explain why income is not perfectly correlated with one's vote (see table 1). Methodologically, it makes sense to assume agents lie in survey studies only if this necessary to account for the actions and facts under consideration, which is not the case here.

voted for left-wing parties, irrespective of their social class and their mobility experience. This gives a strong support to our theory¹¹, which says that in the long-run individual mobility experience has a substantial but completely indirect effect on individual political attitudes: that is, conditioning individual political attitudes on parents' political attitudes cancels almost completely the effect of individual social mobility on voting behaviour depicted in table 1¹².

Also, note that the idea that a common cause leads some agents to support redistribution and to supply less effort is similar to the old view that highly politicized workers do not try to use chances of social ascent as much as workers with less class consciousness¹³ (see Kaelble(1985, p.6)).

Finally, the view that there exists wide and persistent disagreements about the incentive costs of taxation is supported by the strong lack of consensus among economists when they attempt to quantify these costs: everybody agrees that a 90% marginal income tax rate may well discourage labour supply and that a 10% rate leaves room for more taxation, but the consensus is not preserved long if we try to go further. This is hardly surprising since economists face the same basic limitations as the agents described in this paper: the only

¹¹It is hard to reconcile these very high rates of dynastic reproduction with the basic voting patterns of table 1 without a theory giving a common reason why some dynasties vote for more redistribution and at the same time have lower rates of upward mobility.

¹²See also Kelley(1992) for some detailed evidence showing that the effect of social origins is mostly indirect, i.e. goes through the parents' political preference and not the class per se.

¹³This example illustrates that "left-wing dynasties" may very well spend high effort levels for other objectives which are not related to social ascent (such as trade-union activism or teaching).

way to know for sure the optimal redistribution rate would be to try it for a while, and this entails substantial social costs. The difference (hopefully) is that most agents base their assessment on their limited personal experience (so that their eventual beliefs are to a large extent forecastable), whereas scholars perform more sophisticated cognitive processes than those implied by Bayes' rule, and/or have more time to find more information¹⁴.

The rest of this paper is organized as follows: section 2 sets up a simple model of redistribution and learning; section 3 analyzes long-run, steady-state voting patterns and redistribution rates; section 4 shows how sophisticating the collective learning process does not affect the long-run heterogeneity of beliefs but restricts in interesting ways the degree of heterogeneity that one ought to observe in any single country; section 5 attempts to make some outside observer's welfare comparisons of the various steady-states; section 6 gives concluding comments.

Section 2 : A Model of Redistribution and Learning.

In order to highlight the heterogeneity of voting behavior stemming from heterogeneous beliefs, we consider a model of redistribution where different income groups do not a priori have different distributive objectives when they vote over redistributive policies¹⁵. This may

¹⁴Section 5 shows how an outside observer can use our theory and international evidence to make some (limited) progress in assessing these incentive costs.

¹⁵As we repeatedly stress along the paper, a model where voting heterogeneity comes entirely from heterogeneous, well-informed economic interests can not explain the voting

arise either because redistribution is of a pure social-insurance nature (each agent faces equal chances at the beginning of each period), or because all agents share the same distributive-justice principles, although they may have different material interests in redistribution. For reasons dicussed above, we choose to focus on the latter case, at no cost in generality (see below).

We assume a discrete infinite horizon, $t=1,2,\dots$, and we consider an economy made up of a continuum of agents $I=[0;1]$. For conveniency we shall think of each period as a generation, and of each agent as having exactly one offspring each period¹⁶.

During each period each agent can obtain one of two possible pre-tax incomes y_0, y_1 , with $y_1 > y_0 > 0$. We note L_t (resp. $H_t=1-L_t$) the mass of agents born at time t in low-income families (resp. in high-income families)¹⁷.

Agents obtain income y_0 or y_1 depending on luck, how much effort one spent, and social origins (i.e. parents' income). More precisely, the probability that an agents with social origins y_0 (resp. y_1) and with effort supply e obtains income y_1 is given by

$$\text{proba}(y_1 | e, y_0) = \pi_0 + \theta e$$

$$\text{(resp. proba}(y_1 | e, y_1) = \pi_1 + \theta e)$$

patterns of table 1. This does not preclude real-world individual concerns for redistribution to be some complex combination of selfish and "social" values (as long as this is consistent with table 1 and the observed rates of political reproduction).

¹⁶Although nothing would be changed if lifetimes last several periods.

¹⁷ L_t is the mass of agents obtaining income y_0 at time $t-1$.

We assume that $0 < \pi_0 < \pi_1$ to reflect that children from high-income families have access to better opportunities (on average). $\theta > 0$ measures the extent to which individual achievement is responsive to individual effort.

Agents' material welfare is given by

$$U = E(y) - C(e) \text{ }^{18}$$

with $C(e) = e^2/2a$, $a > 0$ ¹⁹ and $E()$ the expectation

After they choose their effort level e and their income shock y_0 or y_1 is realised, agents vote over the redistributive policy τ_{t+1} to be applied next period²⁰. Thus at the time of the vote there are four types of agents: the stable lower-class, noted SL_t (those whose parents' income was y_0 and whose income is also y_0), the downwardly-mobile, noted DM_t (those whose

¹⁸That is, agents maximize their own one-shot utility. We feel comfortable with this zero-discount-factor (infinite time preference) assumption first because the result of long-run heterogeneity of beliefs would hold as long as the discount factor is not too close to 1 (see Rotschild(1974) and Aghion et al.(1991)), and most of all because we view the process of learning about society's mobility parameters (described below) more as an unintended side-product of social experience than as a self-conscious search with optimal experimentation for the sake of future generations (maybe because people do not feel there is anything stationary to learn about mobility). Therefore assuming a zero-discount-factor for this learning process is not contradictory with substantial intergenerational altruism.

¹⁹We assume a to be small enough so that probabilities will always be between 0 and 1. We choose this simple functional form for $C(e)$ for the sake of notational simplicity.

²⁰As to why people go and vote despite their negligible importance, we have nothing original to say. Assume for example that the continuum economy we described so far is in fact a large finite economy with some positive probability of being the decisive voter (the economy must be sufficiently large so that agents' "social" concerns do not show up when choosing effort levels).

parents' income was y_1 and who have gone down to y_0), the upwardly-mobile, noted UM_t (those whose parents' income was y_0 and who have moved up to y_1), and the stable high-income (or middle-class, noted SH_t (those whose parents' income was y_1 and whose income is also y_1)).

We assume that when voting over redistribution these different agents share the same social welfare function. To fix ideas, we assume that they all think that unequal opportunities (i.e. $\pi_0 < \pi_1$) is a bad thing, and that the state should try to correct this as much as possible, i.e. should try to maximize the expected welfare of lower-class children by redistributing income from y_1 to y_0 ²¹. Those readers who feel unhappy with this social objective can replace it by another social welfare function (such as the utilitarian sum of utilities, assuming risk aversion), without changing the substance of what follows (see below).

The important point is that every voter is going to balance the social benefits of equalizing opportunities with the incentive costs of taxation. That is, setting a tax rate τ_{t+1} will lead period- $t+1$'s agents to choose an effort level $e(\tau_{t+1}, \theta)$ maximizing their own expected welfare²²:

²¹Here we assume that the only redistributive policy tool available is pure income redistribution (i.e. tax income at rate τ and redistribute everything in a lump-sum way). Nothing would be modified if we assumed that the state could use public money to act directly on the high-achievement probability π_0 of lower-class children (for example through public schooling).

²²Obviously, there's nothing contradictory between maximizing a "social" objective function when voting and maximizing private welfare when choosing one's effort level: in the latter case, no positive-mass effect is imposed on the aggregate. This is the traditional distinction between private and social values (see, e.g., Arrow(1963, p.18)).

$$e(\tau_{t+1}, \theta) = \text{ArgMax}_{e \geq 0} \theta e(1-\tau_{t+1})(y_1-y_0) - C(e)$$

$$\text{that is: } e(\tau_{t+1}, \theta) = a\theta(1-\tau_{t+1})(y_1-y_0)$$

Taking this into account, the tax rate τ_{t+1} maximizing the expected welfare of lower-class children at period $t+1$ is given by

$$\begin{aligned} \tau_{t+1}(\pi_1-\pi_0, \theta) = \text{ArgMax}_{\tau \geq 0} & (\pi_0 + \theta e(\tau, \theta))(1-\tau)y_1 + (1-\pi_0 - \theta e(\tau, \theta))(1-\tau)y_0 \\ & + \tau [(\pi_0 L_{t+1} + \pi_1 H_{t+1} + \theta e(\tau, \theta))(y_1-y_0) + y_0] - C(e(\tau, \theta)) \end{aligned}$$

that is:

$$\tau_{t+1}(\pi_1-\pi_0, \theta) = H_{t+1}(\pi_1-\pi_0)/a(y_1-y_0)\theta^2$$

Unsurprisingly, the socially-optimal tax rate is an increasing function of $(\pi_1-\pi_0)$ and a decreasing function of θ : the larger the inequality of opportunity $\pi_1-\pi_0$, the more it needs to be corrected, and the higher the income elasticity θ with respect to effort, the more severe the incentive problem²³. Note that these properties do not depend on the particular social welfare function that we chose for illustrative purposes²⁴.

Now, assume that initially agents have different beliefs on society's structural parameters

²³Note also that no public intervention is required if opportunities are equal, i.e. $\pi_1-\pi_0=0$ (this is because we assumed no risk aversion), and that more equalization of opportunities is less costly when the society is richer (i.e. U_{t+1} larger).

²⁴In particular the same properties would hold if one maximizes any (weighted-)utilitarian social welfare function (assuming positive risk aversion, otherwise the optimal utilitarian tax rate is always 0).

(π_0, π_1, θ) . That is, all agree that opportunities are to some extent unequal, but some think that the "deterministic" difference in opportunities $\pi_1 - \pi_0$ is small as compared to the importance θ of individual effort in shaping individual achievements, so that they want very little state intervention so as not to offset individual incentives; whereas some others agree that incentives are a problem, but that overall θ is sufficiently small as compared to $\pi_1 - \pi_0$ (that is, structural and predetermined factors outweigh individual factors) so that the state can play a substantial role in raising revenue to equalize opportunities without that much harm.

The question we want to investigate is the following: assume there is some "true", stationary set of parameters $(\pi_0^*, \pi_1^*, \theta^*)$; what happens in the long-run if agents start with different beliefs on these parameters? what do the long-run voting patterns look like? what role is played by social mobility in this learning and voting process?

To answer these questions, we must first specify how agents learn about society's mobility parameters. Each dynasty $i \in I$ starts at $t=0$ with some prior belief $\mu_{i0}(\cdot)$ defined over the set of all logically possible (π_0, π_1, θ) , chooses an effort level $e_{i0}(\mu_{i0}(\cdot), \tau_0)$ given some (arbitrary) tax rate τ_0 to start with, rationally updates its belief given its income achievement, takes part to the voting process over τ_1 by supporting what one believes to be the socially-optimal policy $\tau_{i1}(\mu_{i1}(\cdot))$ given the posterior belief $\mu_{i1}(\cdot)$ ²⁵, and finally transmits its posterior to its

²⁵We assume here that every voter computes the socially-optimal tax rate as if he thought that everybody shares his beliefs; otherwise he should take into account the fact that others may take higher or lower effort levels than he thought they should. However assuming that he has some beliefs over the distribution of beliefs and that he fully takes into account these effects would not change anything essential; i.e. agents believing in higher θ s would still tend to prefer lower tax rates: if everybody observes a common signal θ of the average beliefs, then the most-preferred tax rate $\tau(\pi_1 - \pi_0, \theta, \theta)$ as a function of one's own

offspring; and so on...

Note that as far as effort-taking and voting is concerned, only the average of the probability measures $\mu_{it}(\cdot)$ are relevant: by linearity,

$$e_{it}(\tau, \mu_{it}) = e(\tau, \theta(\mu_{it}))$$

$$\tau_{it+1}(\mu_{it}) = \tau(\pi_1(\mu_{it}) - \pi_0(\mu_{it}), \theta(\mu_{it}))$$

$$\text{with } \pi_0(\mu_{it}) = \int \pi_0 d\mu_{it}, \pi_1(\mu_{it}) = \int \pi_1 d\mu_{it}, \theta(\mu_{it}) = \int \theta d\mu_{it}$$

This is not the case for the bayesian updating process, however: the entire belief matters. Note also that Bayes' rule puts few restrictions on short-run learning from one's own experience: one's effort level and political attitudes can go in every direction following, say, an upwardly mobile trajectory, depending on how initial beliefs determine the interpretation of the event; we shall see that this ambiguity disappears in the long-run (as arbitrary priors disappear). Finally, note that the voting process is perfectly standard: preferences over tax rates are single-peaked around the most-preferred tax rates $\tau_{it+1}(\mu_{it})$, and the median of these rates is elected and becomes τ_{t+1} ²⁶.

This collective learning process is defined as a sequence of independant single-dynasty learning processes, except that individual experimentation is influenced by some collective

beliefs θ_i is equal to $H_i(\pi_1 - \pi_0) / a(y_1 - y_0)\theta^2 + (1 - \theta_i) / \theta$, which is still decreasing in θ_i for a given θ .

²⁶In lack of a better theory of political parties, we thus assume them to be purely opportunistic. Nothing essential would be changed to individual learning processes had we assumed partisan parties with exogeneous objective functions or beliefs (in particular proposition 3 would still hold).

variable (the redistribution rate τ_t) determined by the cross-section distribution of beliefs. In effect, the learning process we specified above is fully rational if one assumes that each single dynasty to observe only its own economic achievement²⁷ and knows nothing about the rest of the system²⁸. We feel comfortable about this assumption, first because most agents have access to little hard information beyond their immediate family circle (all results would survive if some finite groups of dynasties share common information and experimentation²⁹), and mostly because only under extreme assumptions can one learn substantially more by looking at cross-section aggregates (such as the cross-section distributions of beliefs, income and mobility rates); we discuss these issues in section 4.

Section 3 : Steady-State Political Attitudes.

The first property of this dynamic process of learning and voting is that it converges; that

²⁷One can assume either that each dynasty behaves as a single infinite-horizon bayesian agent (parents "transmit" their posterior to their offspring in the same way as a single agent uses last period's posterior as his new prior), or that each new generation starts its learning life with their own prior and observes the entire mobility history of their family or observes only their parents' posterior and know that they're bayesian (this is equivalent).

²⁸That is, they vote for the redistribution they prefer when given the choice, but they don't where the equilibrium redistribution comes from and don't try to assess its informativeness (we show in section 4 that the latter is very limited, anyway).

²⁹The point is that finite groups of dynasties observing each other's experimentation will tend to experiment the same way (i.e. to take the same effort level) given time preference. A given experimentation would still give more information, but this would only change the size of the pertubations required to remove some wrong belief, without affecting the set of stable beliefs defined next section; as the size of the groups goes to infinity, infinitely small pertubations remove every beliefs except the correct one.

is, in the long-run beliefs about society's mobility parameters and the resulting equilibrium tax rate are stationary³⁰. This is a direct consequence of the martingale convergence theorem.

Proposition 1. For every dynasty $i \in I$, the belief $\mu_{it}(\cdot)$ converges with probability 1 toward some stationary belief $\mu_{i\infty}(\cdot)$ as t goes to ∞ . The equilibrium tax rate τ_i converges toward some tax rate $\tau_{i\infty}$.

Proof. For any given tax rates sequence $(\tau_t)_{t \geq 0}$, the stochastic process $(\mu_{it}(\cdot))_{t \geq 0}$ is defined by a standard, fully-rational process of bayesian updating, and as such has the martingale property (see, e.g., Aghion et al.(1991)). Thus the martingale convergence theorem applies, and the society converges toward some stationary set of beliefs $(\mu_{i\infty}(\cdot))_{i \in I}$. It follows that the equilibrium tax rate, as a continuous function of these beliefs, converges. QED.

Now, the interesting question is whether every dynasty necessarily adopts the same belief in steady-state, and whether the long-run tax rate is necessarily equal to the "true" socially-optimal tax rate. Those readers who are familiar with Rotschild(1974)'s two-armed bandit problem shouldn't be too surprised that the answer to both questions is no³¹: learning

³⁰Obviously, this would not be true if society's mobility parameters are not stationary, which may well be the case in practice. We leave this for future research.

³¹Each effort level is an arm of this continuous-armed bandit problem (note that we make learning much easier by assuming a linear functional form between effort and mobility probabilities; see section 4).

about the relative importance of luck and individual effort in shaping individual achievements is just as complicated as learning the average payoff of a multi-armed bandit's arms. Indeed, knowing how one's social achievement (as a function of social origins) responds to variations in individual (lifetime) effort would require a lot of costly experimentation; namely, several generations would have to "sacrifice" their life by trying to supply no effort at all or to work like mad in order to see what happens to their socio-economic status! Therefore in general not everything is learned in the long-run, and initial beliefs and social mobility trajectories play a key role in shaping actual long-run beliefs and political attitudes.

Although it is impossible to derive analytically the mapping from initial beliefs $(\mu_{i0}(\cdot))_{i \in I}$ to long-run beliefs $(\mu_{i\infty}(\cdot))_{i \in I}$, we are able to say which long-run beliefs $(\mu_{i\infty}(\cdot))_{i \in I}$ are more likely to be observed than others by appealing to some stability criterion. Indeed if we just define an "observable" steady-state as a set of beliefs reproducing itself, then (almost) anything is "observable", and in particular any set of point-beliefs: by definition of bayesian updating, one can never learn what was ruled by the prior. However some of these beliefs are much less likely to be observed than others: typically, beliefs generating expected mobility probabilities that are different from the actual mobility frequencies are very unstable, and conversely. That is, we define $\Delta(\tau)$ as the set of (π_0, π_1, θ) such as the optimal effort level $e(\tau, \theta)$ associated to the point-belief $1_{\tau_0, \tau_1, \theta}$ and the tax rate τ generates a statistical distribution of high and low incomes identical to that expected by an agent with prior $1_{\tau_0, \tau_1, \theta}$; that is, $\Delta(\tau)$ is the set of all (π_0, π_1, θ) such that

$$\pi_0 + \theta e(\tau, \theta) = \pi_0^* + \theta^* e(\tau, \theta)$$

$$\pi_1 + \theta e(\tau, \theta) = \pi_1^* + \theta^* e(\tau, \theta)$$

That is:

$$\Delta(\tau) = ((\pi_0(\theta), \pi_1(\theta) = \pi_1^* - \pi_0^* + \pi_0(\theta), \theta)_{\theta \geq 0})$$

$$\text{with } \pi_0 = \pi_0^* + (\theta^* - \theta)e(\tau, \theta)$$

[Figure 1 represents the locus $\Delta(\tau)$]

Intuitively, if a belief $\mu_{i\omega}(\cdot)$ and a tax rate τ_ω lead to an effort level $e_{i\omega}$ such that the statistical distribution of upwardly- and downwardly-mobile trajectories that does not coincide with that expected by the agent, any small deviation from $\mu_{i\omega}(\cdot)$ in any direction that gets it closer to the true distribution will be recognized by the agent (with some positive probability); conversely, if the "true" statistical distribution and the expected distribution do coincide most perturbations won't be recognized; the point is that there are many beliefs generating actions such as the expected distribution and the statistical distribution coincide (see figure 1): it is difficult to realize that one puts too much weight on effort if one puts at the same time too little weight on predetermined factors. We say that a steady-state is stable if it consists of beliefs that cannot be removed by all small deviations in the direction of either lower or higher mobility than previously expected (see the appendix)³²;

³²One may want to define stability as the property that any individual deviation in some sufficiently close neighborhood of one's stationary belief converges toward the initial stationary belief with probability 1. However this is too demanding: no steady-state is stable according to this definition (not even the "true" belief $1_{(\tau_0^*, \tau_1^*, \theta^*)}$). This comes from topological problems similar to those met by Aghion et al.(1991, pp.636-637): "closeness of

$$\pi'_0 + \theta e(\tau, \theta) = \pi_0 + \theta e(\tau, \theta)$$

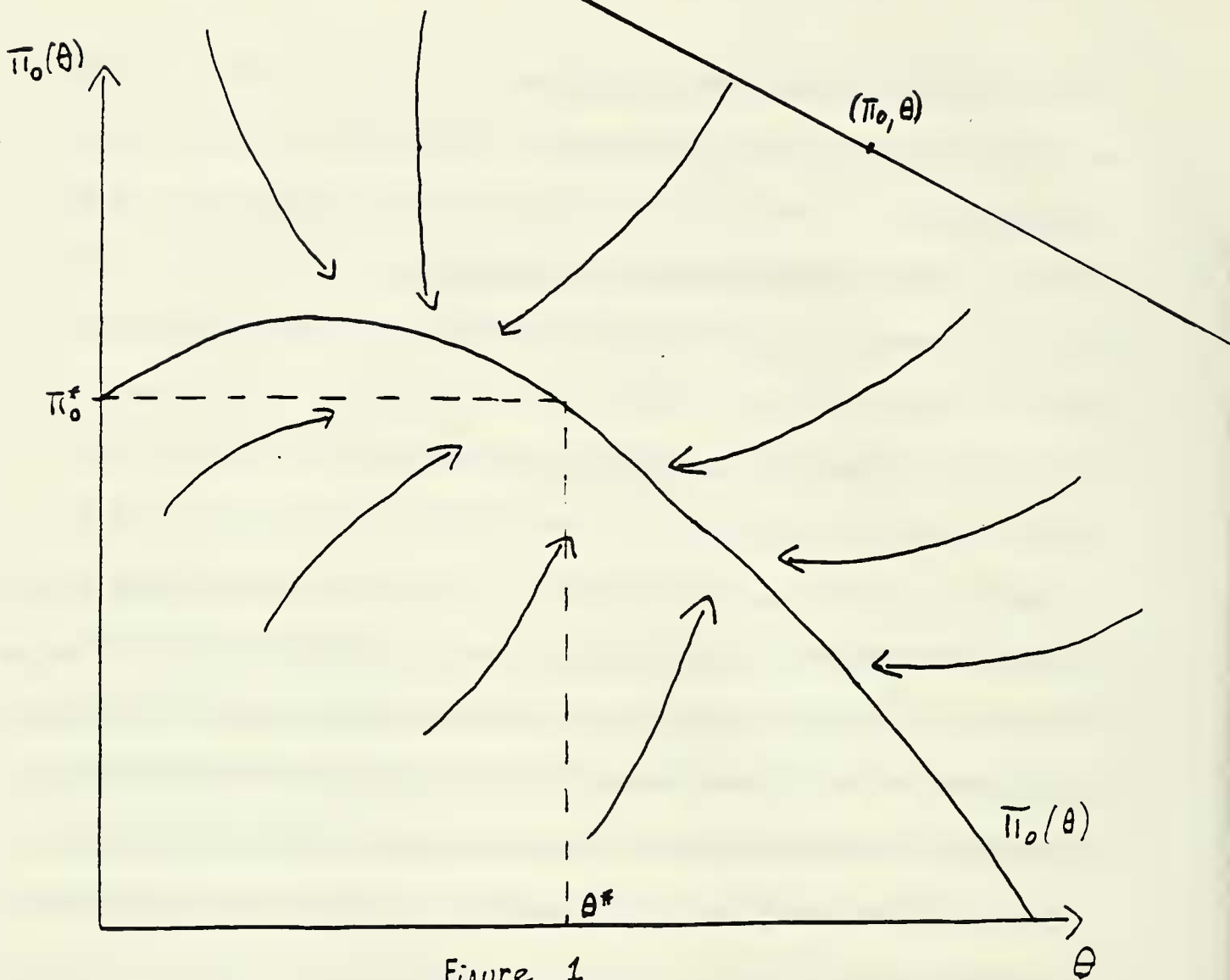


Figure 1

Set of stable beliefs $\Delta(\mathcal{B}) = \{ (\pi_0(\theta), \pi_1(\theta), \theta) \mid \theta \geq 0 \}$

with $\pi_0(\theta) = \pi_0^* + (\theta^* - \theta) e(\tau, \theta)$

$$\pi_1(\theta) = \pi_0(\theta) + \pi_1^* - \pi_0^*$$

this implies that stable, stationary beliefs must have their averages on $\Delta(\tau_o)$.

Proposition 2. $((\mu_{i\in I}(\cdot))_{i\in I}, \tau_o)$ is a stable-steady state iff

$$(1) \forall i \in I, (\pi_0(\mu_{i\in I}), \pi_1(\mu_{i\in I}), \theta(\mu_{i\in I})) \in \Delta(\tau_o)$$

$$(2) \tau_o \text{ is the median of } (\tau_{i\in I})_{i\in I}$$

Proof. see the appendix.

What proposition 2 tells us is that for any stable steady-state all dynasties can be ranked along a one-dimensional scale, namely, their position on the curve $\Delta(\tau_o)$. That is, in the long-run all dynasties believe that the "pre-determined" opportunity difference $\pi_1 - \pi_0$ between lower-class and middle-class children is (on average) $\pi_1^* - \pi_0^*$ (the "true" opportunity difference), but they have different estimates of θ , i.e. of how much individual effort can undo the effects of social rigidities. All dynasties are mobile, so that one can find proponents of all redistributive policies in every income group. But the point is that because the same beliefs lead some dynasties to supply less effort and to support more redistribution, in steady-state there are more left-wing voters among the lower-class (irrespective of who has got the right belief, if any), and the political composition of socially mobile agents is strictly intermediary between that of the stable.

To see that note that those dynasties $i \in I$ who have have converged toward a higher

beliefs" in the topological sense allows for too many deviations.

$\theta = \theta(\mu_{i\omega})$ vote for less redistributive policies ($\tau(\pi_1 - \pi_0, \theta)$ is decreasing with θ) and supply more effort ($e(\tau_{\omega}, \theta)$ is increasing with θ) so that a higher fraction of them $H_{\omega}(\theta)$ has a high-income in steady-state; indeed, $H_{\omega}(\theta)$ is given by the condition that the mass going out of the high-income class is equal to the mass coming in:

$$(\pi_1^* + \theta^* e(\tau_{\omega}, \theta))H(\theta) + (\pi_0^* + \theta^* e(\tau_{\omega}, \theta))L(\theta) = (1 - \pi_1^* - \theta^* e(\tau_{\omega}, \theta))H(\theta)$$

$$\text{that is : } H_{\omega}(\theta) = [1 - 2(\pi_1^* - \pi_0^*)] / [1 - 2\pi_1^* + \pi_0^* - \theta^* e(\tau_{\omega}, \theta)] - 1$$

$$\text{so that } H_{\omega}'(\theta) > 0 \quad (\text{as long as } H_{\omega}(\theta) < 1)$$

It follows that a higher fraction of lower tax rates supporters has a high income. In the same way, lower tax rates supporters have a higher probability of being upwardly-mobile than stable in the lower class, but a lower probability of being upwardly-mobile than stable in the middle class. Indeed the steady-state fractions of θ -dynasties who are upwardly mobile $UM_{\omega}(\theta)$, downwardly-mobile $DM_{\omega}(\theta)$, stable at high-income $SH_{\omega}(\theta)$ and stable at low-income $SL_{\omega}(\theta)$ are given by

$$UM_{\omega}(\theta) = (\pi_0^* + \theta^* e(\tau_{\omega}, \theta))L_{\omega}(\theta)$$

$$DM_{\omega}(\theta) = (1 - \pi_1^* - \theta^* e(\tau_{\omega}, \theta))H_{\omega}(\theta)$$

$$SH_{\omega}(\theta) = (\pi_1^* + \theta^* e(\tau_{\omega}, \theta))H_{\omega}(\theta)$$

$$SL_{\omega}(\theta) = (1 - \pi_0^* - \theta^* e(\tau_{\omega}, \theta))L_{\omega}(\theta)$$

It follows that the fraction of θ -dynasties who are mobile as compared to the fraction of θ -

dynasties who are stable at high-income (resp. low-income) decreases (resp. increases) with θ . Therefore the mobile as a whole have political orientation which are intermediate between those of the stable.

Proposition 3. In any stable steady-state, the voting patterns mimic those presented in table 1. That is, for any two redistributive policies τ, τ' , with $\tau > \tau'$,

$$H_w(\tau, \tau') < L_w(\tau, \tau')$$

$$SH_w(\tau, \tau') < UM_w(\tau, \tau'), DM_w(\tau, \tau') < SL_w(\tau, \tau')$$

where $X(\tau, \tau')$ is the fraction of class X preferring τ to τ'

Proof. Because preferences over tax rates are single-peaked, there exists τ'' , with $\tau > \tau'' > \tau'$, such that dynasties $i \in I$ preferring τ to τ' are those whose most-preferred tax rates $\tau(\theta(\mu_{i\infty}))$ is above τ'' , i.e. those whose $\theta(\mu_{i\infty})$ is below some θ'' . Since the fraction of θ -dynasties $H_w(\theta)$ obtaining a high-income in steady-state increases with θ , the fraction of the high-income class whose $\theta(\mu_{i\infty})$ is below some θ'' is lower than that of the low-income class. Similarly, because $SH_w(\theta)/DM_w(\theta)$ and $SH_w(\theta)/UM_w(\theta)$ increase with θ , $SH_w(\tau, \tau') < UM_w(\tau, \tau')$ and $SH_w(\tau, \tau') < DM_w(\tau, \tau')$, and conversely with SL_w QOED.

Thus in the long-run social origins have an effect on political attitudes because only because they are informative on which type of dynasty one belongs. Prior to convergence however, one cannot completely distinguish between the indirect and the direct effect: many

lower-class agents are in the lower-class because their ideology does not push them to work hard to be promoted, but also their poor economic performance confirms their initial ideology (and conversely for the right-wing ideology).

Section 4 : Robustness of Long-Run Heterogeneity.

We now address the issue of learning by looking at cross-section aggregates and show that this constraints in interesting ways the steady-states without affecting the flavour of the main results (i.e. long-run beliefs heterogeneity and proposition 3)³³.

Consider first how much agents can learn by observing the cross-section distribution of beliefs. If they can observe the exact beliefs $\mu_{i\omega}(\cdot)$ of other dynasties, then in "steady-state" they can infer the true $(\pi_0^*, \pi_1^*, \theta^*)$; this is because the curve $\Delta(\tau_\omega)$ of stationary beliefs depends on the true $(\pi_0^*, \pi_1^*, \theta^*)$: some dynasty i believing in $\mu_{i\omega}(\cdot)$ is ready to accept that some other dynasties have some "wrong" stationary beliefs, but they must be on the curve $\Delta_{(\pi_0(\mu_{i\omega}), \pi_1(\mu_{i\omega}), \theta(\mu_{i\omega}))}(\tau_\omega)$ defined by replacing $(\pi_0^*, \pi_1^*, \theta^*)$ by $(\pi_0(\mu_{i\omega}), \pi_1(\mu_{i\omega}), \theta(\mu_{i\omega}))$ in the definition above³⁴; since these two curves do not coincide dynasty i should realize that $\mu_{i\omega}(\cdot)$ can't be the right belief (see figure 2). Thus assuming that agents observe where the cross-section distribution of beliefs is exactly located in the space of beliefs implies that the

³³We discuss these issues without doing any kind of cost/benefit analysis of information acquisition. The lack of consensus about these issues among those who spend their life studying them (see section 1) suggests that the costs of acquiring the relevant pieces of information are quite high, although the collective benefits may be high.

³⁴Obviously, this is assuming that each agent is able to write and solve the model as we did last section, which is quite demanding. Otherwise, there is not much to be said.

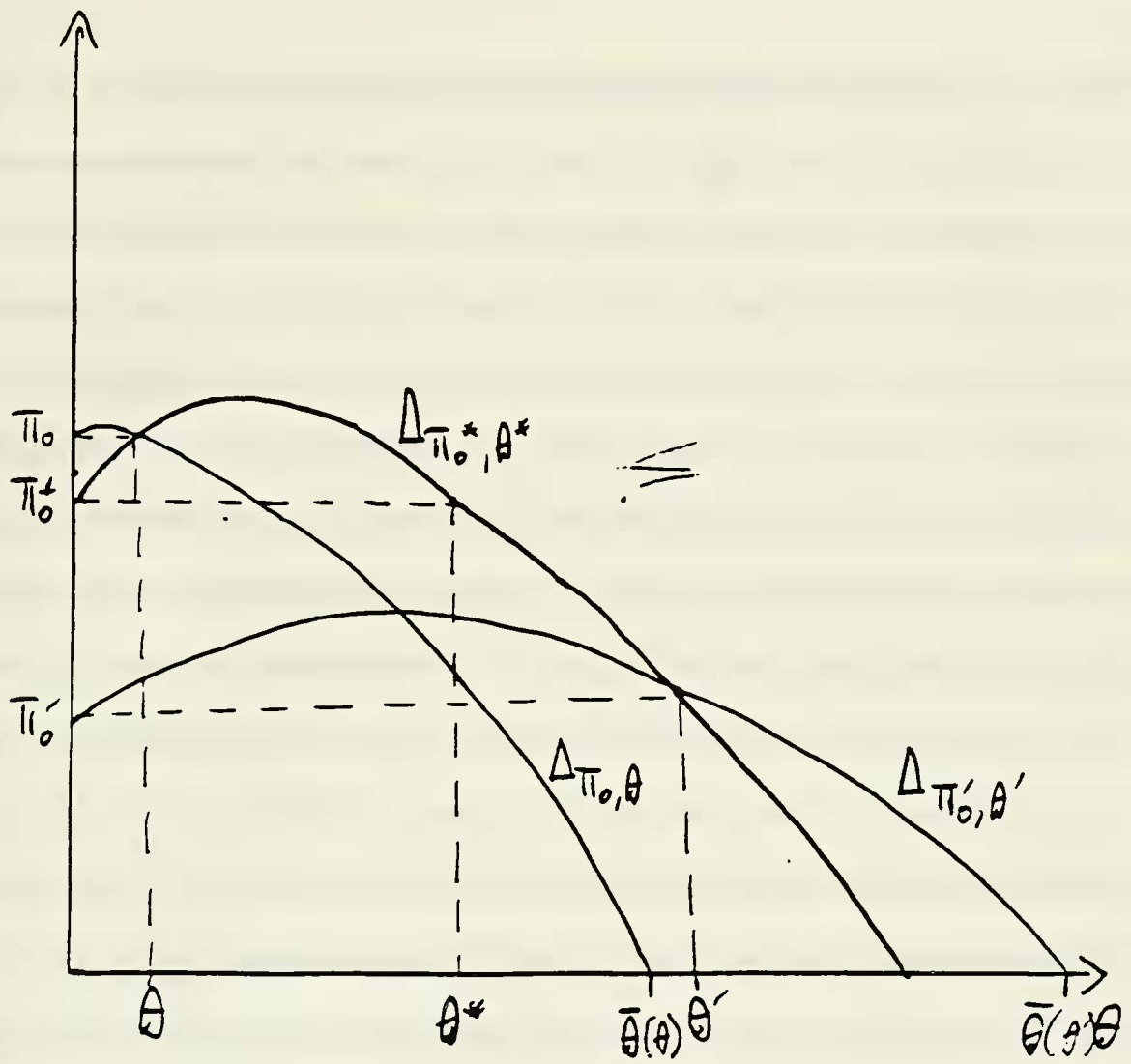


Figure 2

only stable steady-state involves everybody learning the true $(\pi_0^*, \pi_1^*, \theta^*)$.

However this inference relies on the unrealistic assumption that agents can observe the exact probability measure $\mu_{i\omega}(\cdot)$ of other agents; in practice, the only "material" expression of a belief is an effort level $e_{i\omega}(\mu_{i\omega})$ and a most-preferred tax rate $\tau_{i\omega}(\mu_{i\omega})$, and an agent can observe at most the cross-section distribution of these choice variables. These are very insufficient statistics for the entire beliefs, and one can easily see from figure 2 that a rational bayesian updater can't learn much by observing these cross-section distributions: the only steady-states that are ruled out are those involving extreme left-wing agents and extreme right-wing agents at the same time. This is because self-consistent, fully-rational left-wing dynasties believing in a low $\theta(\mu_{i\omega})$ believe that the maximum steady-state "mistake" is $\theta_{\max}(\theta(\mu_{i\omega}))$, which is lower than the "true maximum mistake" $\theta_{\max}(\theta^*)$ (see figure 2); that is, if these dynasties observe too right-wing dynasties, they must infer that something is going wrong and revise these beliefs³⁵. Therefore any steady-state must be such that the most left-wing beliefs $\mu_{i\omega}$ and the most right-wing beliefs are mutually compatible, i.e. are such that $\theta_{\max}(\theta(\mu_{i\omega})) \geq \theta(\mu_{i\omega})$, which limits slightly the set of steady-states defined in proposition 2.

The other inference that fully-rational agents could make by looking at the cross-section distribution of beliefs (as expressed by, say, the most-preferred tax rates) would be to think about what the initial distribution of priors was when this all process started long ago in the past, to compute what steady-state distribution of beliefs this implies as a function of

³⁵Note that this reasoning does not hold for extreme right-wing dynasties, who can move on with their extreme beliefs without ever worrying about the extreme left-wing people in the other corner!

the true $(\pi_0^*, \pi_1^*, \theta^*)$, and then to try to invert this mapping. This strikes us as an extreme implication of bayesian rationality (in the history of inequality and redistribution the notion of a time 0 is rather elusive, and for sure real-world agents don't perform this thought experiment), but in any case this would not result into adequate learning. The reason is that this mapping has no reason to be invertible: the dimension of the set of long-run distribution of observed actions has no reason in general to be larger than the dimension of the set of unknown parameters³⁶, so that observing the former gives limited information on the latter. Therefore such an inference process could reduce somewhat the set of "admissible" steady states³⁷ but would not change the basic result of long-run heterogeneity and voting patterns³⁸.

Consider now what agents could learn by observing both the cross-section distribution of actions and the cross-section income distribution and mobility rates. Assume first that they observe the steady-state distribution of most-preferred tax-rates $(\tau_{i\omega}(\mu_{i\omega}))_{i \in I}$ (or of effort

³⁶In our model, the set of all logically possible distributions of actions does have higher dimensionality than the set of possible parameters, but this is an artifact of the two-income modelling and of the linearity of the mobility process: in general the set of all parameters required to describe how the entire mobility matrix responds to effort is very likely to be much higher-dimensional than the observable part of the distribution of actions.

³⁷Typically, this could again rule out steady-states with too extreme beliefs on both sides, since this inference process gives the same information (if any) to everybody which implies a relative "leveling" of all beliefs.

³⁸This confirms the finding by Smith(1991) that this kind of inference from observing others' actions can result into adequate learning only under extreme assumptions: although Smith makes the adequate dimensionality assumptions to avoid invertibility problems, adequate learning occurs only if the distribution of priors contains unboundedly informative priors (but convergence toward a common belief is guaranteed).

levels $(e_{i\omega}(\mu_{i\omega}))_{i \in I}$ and the steady-state distribution of income $(L_\omega, H_\omega = 1 - L_\omega)$. By observing the distribution of actions they can compute the average effort level $e_\omega = \int e_{i\omega}(\mu_{i\omega}) di$, and combined with the knowledge of H_ω this gives to every agent the true knowledge of the probability of obtaining a high income with effort e_ω : the true $(\pi_0^*, \pi_1^*, \theta^*)$ must verify $\pi_0^* L_\omega + \pi_1^* H_\omega + \theta^* e_\omega = H_\omega$. Combined with one's own beliefs $\mu_{i\omega}$ this allows every agent to infer the true $(\pi_0^*, \pi_1^*, \theta^*)$.

Again, such a successful inference relies on very extreme informational assumptions: that is, the aggregate distribution of effort levels per se is certainly not observable³⁹, and the distribution of most-preferred tax rates is observable only to the extent that they materialize into actual support for some political party. Assume for example that one can only observe the eventual equilibrium tax rate τ_ω , and that the latter is known to be the median of the distribution of most-preferred tax rates⁴⁰. Then going from this observation to the knowledge of the average effort level e_ω is literally impossible: distributions with a fixed mean can have all sorts of average, plus the relation between effort levels and most-preferred tax rates is non-linear.

Assuming that one can observe the relative popularity of two exogeneously-given political parties or the actual mobility rates would not alter the basic message: by observing aggregate characteristics of the collective learning process one can get at most an approximate knowledge of the "aggregate experiment"; this can possibly allow some limited

³⁹Others' effort levels are possibly observable for small, finite groups of agents, which would not change anything to the set of stable steady-states (see footnote 29).

⁴⁰Which in practice is quite speculative.

inference, but in any case extreme assumptions are required to obtain adequate learning.

In sum, even assuming that agents have very sophisticated cognitive ability (they know the right model and are able to compute its dynamic properties), no realistic assumption regarding what agents can observe affects significantly the analysis of the collective learning process: inferring information from looking at cross-section aggregates can typically rule out steady-states with too extreme beliefs and therefore implies a relative "homogenisation" of steady-states⁴¹, but cannot remove the long-run heterogeneity of beliefs.

Section 5 : Some Welfare Analysis.

Now consider an outside observer knowing the model and looking at the pieces of international evidence that we have on inequality, mobility and redistribution in western democracies. Assume also that this outside observer is ready to assume that these countries have the same structural parameters $(\pi_0^*, \pi_1^*, \theta^*)$. The first piece of international evidence is that important and fairly stable differences in levels of redistribution are being observed:

⁴¹The observation of a common signal always has a "levelling effect" on a set of heterogeneous beliefs. Note that there is another reason why too different beliefs cannot sustain in steady-state, which operates through the influence of the equilibrium redistribution on individual experimentations: for example in a country with a tradition of a low τ supplying little effort is more costly (for a given left-wing belief) and therefore it is harder to learn about a possible low θ^* . We believe these effects explain partly the political homogeneity of countries like the US and why the spectrum of political attitudes overlaps so little between both sides of the Atlantic.

typically, there tends to be much less redistributive transfers in the United States than in Western Europe and especially Scandinavia⁴². From this one can infer that these countries are in different steady-state equilibria of the model (this is confirmed by the observation that working hours, i.e. some limited signal of effort, tend to be longer in the US). Of course, if the outside observer looking at these countries knows the true parameters, he can easily say which country redistributes too much, which country works too much, which agents have a "wrong" ideology, and so on: he knows that the "truly optimal" rate of redistribution τ^* , effort level e^* and GNP $L^*y_0 + H^*y_1$ are given by the true parameters $(\pi_0^*, \pi_1^*, \theta^*)$.

But if the outside observer does not know a priori which beliefs are the right ones (just as us), what can he say if he wants to compare the actual welfare of these different dynasties and countries? The answer may first seem to be: not much. Indeed, one can find steady-state where the agents spending the highest amount of effort are in fact not working enough (given the true returns to effort), and others where the agents spending the lowest amount of effort work too much. Maybe there is too much redistribution in the US, and maybe there is too little in Sweden.

In a desperate need to refine his beliefs, the observer may compare the GNPs of these different countries: the theory predicts that a country with less redistribution should have a higher GNP (whatever the true parameters), and that this should be all the more so if the incentive problem is more severe (that is, if the true social optimum is relatively little

⁴²It is hard to give a global quantification of this multi-dimensional phenomenon. Mueller(1989,p.326) presents some data showing that the size of transfers as a fraction of GNP is twice as large in Western Europe than in the US.

redistribution). Here the evidence is not very conclusive: EC countries tend to have a somewhat lower GNP/capita than the US, but this is not so for Scandinavia. Coming down to less and less secure grounds, the observer may want to compare mobility rates: the theory again predicts that countries with less redistributive taxation should have higher mobility rates, and again this should show up particularly strongly if individual incentives play the key role postulated by these countries. The striking observation here is that all quantitative studies that have tried to compare mobility rates across developed countries have concluded that these rates were amazingly similar (see the references given in section 1). The observer may choose to conclude that since the rigid, redistribution-intensive societies of Western Europe are as mobile as the US, there is little reason to believe in such a strong need to preserve individual incentives. This is a very unsecure inference process, but this may be the best one can do to refine arbitrary priors, and we believe this is the kind of international comparisons on which a number of observers "decide" on which side of the Atlantic are we closer to the social optimum.

In theory, one can say more than that by looking in more details at the class composition of the electorates supporting different redistributive policies (say, different parties). For example, if there is a lot of class polarisation (i.e. very high partisan voting in each social class), this suggests that (at least) one class is very far from its socially-optimal welfare level. In the same way, very different most-preferred policies (i.e. main political parties advocating very different rates) suggest that (at least) some dynasties have got it all wrong. Assume now we observe very different policy proposals, but very little class polarisation. This suggests that very different effort levels do not have a major effect on individual

achievements, and therefore that the truly socially-optimal policy involves a lot of redistribution and that those working the most should slow down. Similarly, substantial class polarisation around comparable policy proposals indicate individual factors are the key to success and that the social optimum involves little redistribution. This analysis of class polarization of electorates vs polarization of the political spectrum can also be conducted at the cross-country level. One would have to look at this data in more details, but there does not seem any striking dissimilarity across western countries from which information could be inferred. In any case, these are again very approximate ways to infer some information, but these may be the best ones available given what we want to learn.

Section 6 : Concluding Comments.

This paper has two main objectives. First, providing some theoretical foundations to understand better the political economy of redistribution and particularly some important stylised facts concerning the effect of social mobility on political attitudes toward redistribution (namely, the fact that voters with identical incomes but different social origins vote differently). This gives a richer picture of redistributive politics than the standard median-voter model (which cannot account for these stylised facts). We believe that our theory also provides a tractable framework to analyze the fluctuations of redistributive politics, e.g. one that can be used to look at the effects of changes in the pre-tax distribution on redistributive policies (for example, we have not analyzed how shocks to fundamentals

determine transitions between steady-states).

Next, this paper suggests that instead of always looking at politics as a game of conflicting-interests aggregation, it may be sometime valuable to consider that the main difference between voters is not their differing interests and objective functions but rather the information and ideas about policies that they have been exposed to during their social life. Not only the majority rule is ill-suited to aggregate conflicting interests (see, e.g., majority cycles), but differing beliefs and ideas about government intervention in the economy are pervasive (not only among economists). The point is that although people can have different beliefs about the best-possible policy, these beliefs are not arbitrary: agents are naturally exposed to different pieces of information depending on their economic position. We hope this general approach can be tractable and rewarding enough to solve interesting political-economy questions in the future⁴³.

Appendix.

Proof of proposition 2.

We first define formally the notion of stability that we're using (we restrict formal notations to the case of single-point beliefs (Dirac measures), but everything can be readily extended to non-single-point stationary beliefs by replacing them by their averages).

⁴³For example, consider the unemployment model with voting over firing costs of Saint-Paul(1993): unlike in Saint-Paul's theory, it may well be that there exists some socially-optimal, positive firing costs depending on how much employers internalize the human-capital social costs of firing; in such a case, it may be reasonable to expect that various employment histories lead to various informational exposures regarding employers' "excessive" propensity to fire, leading to different political attitudes and possibly important positive and normative implications.

Consider some stationary belief $l_{(\tau_0, \pi_1, \theta)}$ and some stationary tax rate τ , leading to some effort level $e(\theta, \tau)$ and to some expected mobility probabilities $(\pi_0 + \theta e(\tau, \theta), \pi_1 + \theta e(\tau, \theta))$. Assume (π_0, π_1, θ) is not on $\Delta(\tau)$. Then we prove that any (arbitrarily) small perturbation in the direction of $\Delta(\tau)$ will remove this belief with positive probability (see figure 1). In that sense only beliefs located on $\Delta(\tau)$ can form stable steady-states, in the sense that constantly pertubated beliefs always tend to come back there.

Assume for example that (π_0, π_1, θ) is above $\Delta(\tau)$ (see figure 1), i.e. that $\pi_0 + \theta e(\tau, \theta) > \pi_0^* + \theta^* e(\tau, \theta)$ (the expected upward mobility probability is higher than the true probability). Consider any other parameters $(\pi_0', \pi_1', \theta')$ predicting a lower probability of upward mobility than (π_0, π_1, θ) (for example the parameters of $\Delta(\tau)$; see figure 1), i.e. such that

$$\pi_0 + \theta e(\tau, \theta) > \pi_0' + \theta' e(\tau, \theta)$$

Consider the dynastic learning process starting with beliefs $\mu_{i0} = (1 - z_0) l_{(\pi_0, \pi_1, \theta)} + z_0 l_{(\pi_0', \pi_1', \theta)}$, where z_0 is some arbitrarily small number. We prove that the long-run beliefs $\mu_{i\infty} = l_{(\pi_0, \pi_1, \theta)}$ with probability 0. We note $e(z) = e(\tau, (1 - z)\theta + z\theta')$. Then transition rules for $(z_t)_{t \geq 0}$ are given by

$$z_{t+1}^+ = (\pi_0' + \theta' e(z_t)) / [(1 - z_t)(\pi_0 + \theta e(z_t)) + z_t(\pi_0' + \theta' e(z_t))] z_t$$

if dynasty i observes upward mobility

$$z_{t+1}^- = (1 - \pi_0' - \theta' e(z_t)) / [(1 - z_t)(1 - \pi_0 - \theta e(z_t)) + z_t(1 - \pi_0' - \theta' e(z_t))] z_t$$

if dynasty i observes stability at low income

Since $\pi_0 + \theta e(\tau, \theta) > \pi_0' + \theta' e(\tau, \theta)$, $z_{t+1}^- < z_t < z_{t+1}^+$. For z_t sufficiently small, $(1 - z_t)(\pi_0 + \theta e(z_t)) + z_t(\pi_0' + \theta' e(z_t)) < \pi_0^* + \theta^* e(z_t)$ (by continuity). It follows that for z_t sufficiently small, $Ez_{t+1} = (\pi_0^* + \theta^* e(z_t))z_{t+1}^+ + (1 - \pi_0^* - \theta^* e(z_t))z_{t+1}^- > z_t$ (Ez_{t+1} is the "true" expectation of z_{t+1} , as opposed to $E(z_{t+1} | z_t)$ which by definition is equal to z_t). Finally, $Ez_{t+1} > z_t$ for z_t sufficiently small implies that Ez_t converges to a strictly positive limit, and therefore that z_t cannot converge to 0 with probability 1. One can prove in the same way the instability of any (π_0, π_1, θ) located below $\Delta(\tau)$. CQFD.

References.

- Abramson, P.(1973), "Intergenerational Social Mobility and Partisan Preference in Britain and Italy", *Comparative Political Studies* 6,2.
- Aghion, P., P. Bolton, C. Harris and P. Julien(1991), "Optimal Learning by Experimentation", *Review of Economic Studies*.
- Alesina, A., and R. Perotti(1993), "Income Distribution and Investment", mimeo Harvard.

- Arrow, K.(1963), "Social Choice and Individual Values", Wiley.
- Boy, D.(1980), "Système Politique et Mobilité Sociale", *Revue Française de Science Politique* 30, 925-958.
- Cherkaoui, M.(1992), "Mobilité", in "Traité de Sociologie", R. Boudon Ed. , Presses Universitaires de France.
- Erikson, R. and J. Goldthorpe(1985), "Are American Rates of Social Mobility Exceptionally High?", *European Sociological Review* 1, 1-22.
- Erikson, R. and J. Goldthorpe(1992), "The Constant Flux: A Study of Class Mobility in Industrial Societies", Clarendon Press.
- Kaelble, H.(1985), "Social Mobility in the 19th and 20th Centuries: Europe and America in Comparative Perspective", St-Martin's Press.
- Kelley, J.(1992), "Social Mobility and Politics in the Anglo-Saxon Democracies", in Turner(1992).
- Kluegel, J., and E. Smith(1986), "Beliefs About Inequality", NY:Aldine de Gruyter.
- Lipset, S.M.(1966), "Elections: the Expression of the Democratic Class Struggle", in Bendix-Lipset Ed., "Class, Status and Power", Free Press.
- Lipset, S.M.(1977), "Why no socialism in the US", in "Sources of Contemporary radicalism", Bialer-Blutzar Ed., Westview Press.
- Lipset, S.M.(1992), "Foreword: the Political Consequences of Social Mobility", in Turner(1992).
- Lipset, S.M., and R. Bendix(1959), "Social Mobility in Industrial Society", Univ. of California Press.

- Marx, K.(1852), "Der Achtzehnte Brumaire des Louis Napoleon".
- Miller, D.(1992), "Distributive Justice: What the People Think", *Ethics* 102, 555-593.
- Mueller, D.(1989), "Public Choice II", Cambridge University Press.
- Perotti, R.(1992), "Income distribution, Politics and Growth", *American Economic Review* (Papers and Proceedings).
- Petersen, W.(1953), "Is America still the land of opportunity?", *Commentary* 17, 477-484.
- Piketty, T.(1993), "Dynamic Voting Equilibrium, Political Conservatism and Income Redistribution", Working Paper DELTA.
- Rothschild, M.(1974), "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory* 9, 185-202.
- Rytina, J., W. Form and J. Pease(1970), "Income and Stratification Ideology: Beliefs About the American Opportunity Structure", *American Journal of Sociology* 75, 703-716.
- Smith, L.(1991), "Error Persistence, and Experiential Vs Observational Learning", mimeo MIT.
- Smobart, W.(1906), "Warum gibt es in den Vereinigten Staaten keinen Sozialismus?".
- Thomson K.(1971), "A Cross-National Analysis of Intergenerational Social Mobility and Political Orientation", *Comparative Political Studies* 4, 3-20.
- Tocqueville, A. de(1835), "De la Démocratie en Amérique"
- Turner, F., Ed.(1992), "Social Mobility and Political Attitudes: Comparative Perspectives", Transaction Publishers.

MIT LIBRARIES



3 9080 00846624 2

Date Due

JAN. 03 1995

MAR. 25 1997

MAY 31 2001

OCT 31 2006

Lib-26-67

