

Thema

UMR 8184

Théorie Économique, Modélisation et Applications

Thema Working Paper n° 2009-01
Université de Cergy-Pontoise, France

Equality of opportunity and luck:
Definitions and testable conditions, with an
application to income in France.

Arnaud LEFRANC, Nicolas PISTOLESI, Alain TRANNOY

March 2009



Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France*

Arnaud LEFRANC[†] Nicolas PISTOLESI[‡] Alain TRANNOY[§]

Revised version : March 2009
First version : July 2006

Abstract

We offer a model of equality of opportunity that encompasses different conceptions expressed in the public and philosophical debates. In addition to *circumstances* whose effect on outcome should be compensated and *effort* which represents a legitimate source of inequality, we introduce a third factor, *luck*, that captures the random factors whose impact on outcome should be *even-handed* for equality of opportunity to be satisfied. Then, we analyse how the various definitions of equality of opportunity can be empirically identified, given data limitations and provide testable conditions. Definitions and conditions resort to standard stochastic dominance tools. Lastly, we develop an empirical analysis of equality of opportunity for income acquisition in France over the period 1979-2000 which reveals that the degree of inequality of opportunity tends to decrease and that the degree of risk of income distributions, conditional on social origin, appears very similar across all groups of social origins.

JEL Codes: D63, J62, C14

Keywords: Equality of opportunity, Luck, Stochastic dominance, Income distribution.

*This paper is part of a research program supported by the French Commissariat au Plan. We are grateful to Gérard Forgeot from the French national statistical Agency (INSEE) for the access to the data BdF 2000. For useful comments, we wish to thank François Bourguignon, Christine Chwaszcza, Russell Davidson, Jean-Yves Duclos, Marc Fleurbaey, Francisco Ferreira, Dirk Van de Gaer, Nicolas Gravel, John Roemer, Erik Schokkaert and participants in seminars at the French ministry of the economy (Fourgeaud seminar), the University of Essex (ISER), the University of Oxford, the University of Ghent (Public economics seminar), the University of Verona (Canazei Winter School) and the European University Institute (RSCAS).

[†]Robert Schuman Center for Advanced Studies, European University Institute and THEMA, Université de Cergy-Pontoise. Email: Arnaud.Lefranc@iue.it

[‡]THEMA, Université de Cergy-Pontoise. Email: nicolas.pistolesi@eco.u-cergy.fr

[§]EHESS, GREQAM-IDEF. Email: alain.trannoy@ehess.cnrs-mrs.fr

“The ownership of personal or material productive capacity is based upon a complex mixture of inheritance, luck and effort, probably in that order of relative importance”

Frank Knight, *The Ethics of Competition*, 1935

1 Introduction

Most economic analysis of inequality, theoretical and empirical, rely on the assumption that equality of individual outcomes (*e.g.* welfare, income, health) is *per se* a desirable social objective. This is sometimes criticized for standing at odd with both public perceptions of inequalities and some developments in modern theories of justice. According to this criticism, a distinction must be drawn between morally or socially justified and unjustified inequalities. This has led egalitarian philosophers such as Rawls (1971), Dworkin (1981a, 1981b), Sen (1985), Cohen (1989) or Arneson (1989, 1990) to claim that distributive justice does not entail the equality of individual outcomes but only requires that individuals face equal *opportunities* for outcome. Despite the growing political audience of this view, few economic analysis have tried to assess the extent to which equality of opportunity is empirically satisfied.¹ Two major issues are likely to account for this state of affairs. First, how should equality of opportunity be characterized? In fact, no consensus has been reached, neither in the philosophical nor in the public debates, regarding how opportunities should be defined and in what sense they should be considered equal. In this paper we offer a model of equality of opportunity that encompasses several conceptions expressed in these debates. Second, how can equality of opportunity be empirically assessed? This requires that the determinants of individual outcomes be taken into account. However, these determinants are never fully observable. Hence, we analyze how the various conceptions can be empirically identified, given data limitations, and provide testable conditions for equality of opportunity. Lastly, we develop an empirical implementation of these conditions and examine the extent to which equality of opportunity is achieved in the distribution of income in France.

One important implication of the equal-opportunity view is that judgements about equality must take into account the determinants of individual outcomes. At least two sets of factors must be distinguished : on the one hand, factors that are considered a legitimate source of inequality; on the other hand, factors that do not appear as socially or morally acceptable.

Following the terminology introduced in Roemer (1998), we refer to the former determinants

¹Roemer et al. (2003), O’Neill et al. (2000), Checchi et al. (1999), Benabou & Ok (2001), Bourguignon et al. (2007), Goux and Maurin (2003), Alesina and La Ferrara (2005) and Checchi & Peragine (2005) who analyze equality of opportunity for income and Schuetz et al. (2005) who examine educational opportunities are some of the exceptions.

as *effort* and to the latter as *circumstances*. As most authors would agree, the principle of equality of opportunity essentially requires, that, given individual effort, circumstances do not affect individual prospects for outcome, or to paraphrase Rawls (1971, p.63), that individual with similar effort face “*the same prospects of success regardless of their initial place in the social system*”. What factors should count as effort or circumstances is of course a crucial point from a normative perspective. For several authors, individual responsibility should be the relevant criterion and all factors beyond the realm of individual responsibility should count as circumstances. There remains, however considerable debates on this issue.

A prominent view in these debates is the one expressed by John Roemer in a series of contributions.² ³ It claims that the definition of circumstances is a matter of political choice. Furthermore, once circumstances have been defined “*by society*”⁴, remaining differences in individual outcomes should be considered the result of effort. Hence, the distinction between circumstances and effort turns into a dichotomic partitioning of the determinants of outcome. As a consequence, requiring that, for a given level of effort, circumstances do not affect individual prospects for outcome, implies that individuals with similar effort should have *equal* outcomes.

This dichotomic approach lies at the heart of most economic analysis of equality of opportunity. However, it does not fully account for the diversity of the determinants of outcome and leads to a specific conception of equality of opportunity. Assuming that society has agreed on a given set of circumstances does not imply that the remaining determinants will reflect individual responsible choice and should be treated as effort. In this respect, international attitudes surveys, such as the one summarized in Figure 1, reveal two noteworthy differences across countries. Consider that “social injustice” captures inequalities arising from circumstances, as defined “by society”. First, countries differ in their propensity to consider that bad economic outcomes reflect social injustice, which indicates that the definition of circumstances may vary across societies. Second, and more importantly, the figure also suggests that countries differ in their belief in the role of effort in shaping individual outcomes, over and beyond the influence of circumstances.⁵ The assumption that the determinants of outcomes excluded from socially defined circumstances relate to individual effort provides a good approximation of US average beliefs. It does not however correspond to the social perception in many European countries,

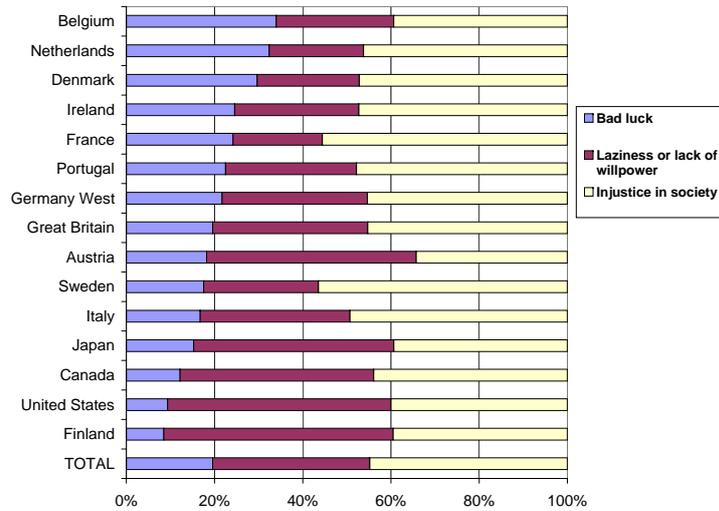
²For a theoretical discussion, see Roemer (1993, 1998) and for empirical applications Betts & Roemer (2006), Roemer et al. (2003) and Dardanoni et al. (2005).

³See Fleurbaey & Maniquet (2007) for a thorough discussion of alternative perspectives and related issues.

⁴Roemer (1993, p.149)

⁵For more detailed evidence, see among others Marshall et al. (1999), Corneo & Gruner (2002), Alesina & Glaeser (2004) and Alesina & Angeletos (2005).

Figure 1: Beliefs in the role of luck, effort and social injustice in bad economic outcomes



Source : World Values Survey (1990). Answers to the question : "Why are there people living in need ?". Authors' computations excluding the following answers : It is an inevitable part of modern progress; None of theses; Don't know.

which emphasizes the role of luck in shaping individual success.

Our purpose is to build a model of equality of opportunity flexible enough to encompass this diversity of perceptions. This requires to distinguish three *generic* determinants of individual outcomes : circumstances, effort and luck. As described in the philosophical literature, justice does not necessarily command that the impact of every kind of luck be nullified. In some cases, luck may appear as a fair source of inequality provided that it is decorrelated from circumstances, in short, even-handed. Consequently, it is too strong a requirement to define equality of opportunity as a situation where individuals with similar effort reach equal outcomes. What equality of opportunity requires is that, given effort, no one faces more favorable outcome prospects, as a result of luck, for reasons related to differential circumstances.

The first contribution of this paper is to offer a characterization of equality of opportunity consistent with this view. Given effort, the outcome prospects of an individual are summarized by the outcome distribution conditional on her circumstances. Our characterization rests on the idea that equality of opportunity prevails when the conditional outcome distribution attached to all possible circumstances cannot unanimously be ranked in terms of well-being, using the tools of stochastic dominance (first and second order). The introduction of luck as a determinant of outcome fully legitimates the use of stochastic dominance instruments, and specifically of second order. Because choosing among the outcome prospects falls down into decision-making problems under risk. If a decision maker is unable to choose among the different prospects,

then EOP prevails. This leads us to distinguish several definitions of EOP.

The empirical implementation of these definitions of equality of opportunity would be straightforward if circumstances and effort were observable. However, in practice, this condition may not be easily met. In most data sets, not all the relevant aspects of individual effort can be measured and only a subset of the relevant circumstances can be observable. We discuss the consequences of these limitations for the evaluation of equality of opportunity. The second contribution of the paper is to show that, conditional on further distributional assumptions, it is still possible in some cases to provide testable conditions for equality of opportunity when effort and circumstances are not fully observed.

We then develop an empirical analysis of equality of opportunity for income acquisition in France, using household surveys over the period 1979-2000. In this application, we assume that circumstances are defined by individual social background, measured by father's occupation and we compare income distributions conditional on social origin. Our analysis of these income distributions relies on non-parametric tests of stochastic dominance developed by Davidson and Duclos (2000).

The rest of the paper is organized as follows. Section 2 discusses our characterization of equality of opportunity. We first review the various conceptions of equality of opportunity that have been discussed in recent philosophical debates and then develop a comprehensive model that accommodates these various conceptions. In section 3, we discuss the identification of equality of opportunity when the relevant determinants of outcome are only partially observable. In section 4, we develop an empirical analysis of equality of opportunity for income in France.

2 Equality of opportunity : Definitions

Our characterization of equality of opportunity rests on the following representation of the determinants of individual outcomes. Individual success or failure is driven by a mixture of deterministic and random factors. Unless explicitly specified, we assume full observation of these determinants and their impact. The random factors can be described as real-valued random variables. A positive value is generally described as good luck and a negative one as bad luck. Why is it important to distinguish random from deterministic factors? From the point of view of equality, three kinds of operation can be described. For both factors, whether deterministic or random, we can require either to erase or on the opposite to fully respect the effect of the factor on the outcome. The rationale for these operations is discussed in length in the equality of opportunity literature about effort and circumstances. For a random factor, one more operation

may be defined: to neutralize its correlation with a given set of deterministic factors. Consider for instance the random factors associated with job search. Now take individual ethnicity into consideration, as a deterministic factor. Equality of opportunity would probably require that the outcome of the job search process be ethnically neutral, i.e. distributed independently of ethnic origin. It may not entail that job search luck should be fully neutralized. In the next paragraphs, we review on what normative principles we should decide whether a given random factor should be neutralized, respected or decorrelated.

2.1 Luck and equality of opportunity : a brief review

In the philosophical debates on equality of opportunity, the concept of luck refers to situations where individual control, choice or moral responsibility bears no relationship to the occurrence of outcomes. As we now discuss, this broad concept includes the notions of circumstances and luck that we previously referred to.^{6 7}

2.1.1 Varieties of luck

The debates about equality of opportunity have singled out at least four ideal-type notions of luck that can be illustrated by simple empirical examples.

First, consider two equally talented and motivated individuals whose outcome differ only because of differences in their family's social connections. In this situation individual actions and their results are pre-determined by antecedent factors (family and social origin). This illustrates the idea of *social lottery* developed by Rawls. It is most probably the first candidate to be considered as a circumstance.⁸ We propose to call it *social background luck*.

Second, consider two fraternal twins whose outcome differ only because one of them genetically inherited a special talent. As in the previous example, the determinant of differential success, talent, lies beyond the realm of individual choice or control. One important difference with the previous form of luck is that a specific individual talent can be seen as constitutive of the individual, in the sense that it defines what person she is. This second example illustrates the notion of *constitutive luck*, or Rawls' idea of a *natural lottery*. It includes genetically inherited factors and we therefore propose to call it *genetic luck*.

Third consider two individuals with similar talent and social background. Their outcomes

⁶In welfare economics, the analysis of social situations in the presence of risk has revealed a conflict between the *ex ante* and *ex post* approaches. See Diamond (1967), Hammond (1981), Broome (1991), Ben Porath et al. (1997), Gajdos and Maurin (2004) and Fleurbaey (2006). This perspective is however essentially absent from philosophical debates on equality of opportunity.

⁷See Lippert-Rasmussen (2005) for a discussion of the relationship between luck and distributive justice.

⁸See for instance the discussion in Dardanoni, Fields, Roemer & Sanchez Puerta (2005).

differ as a result of a lottery they could not escape. For instance, as a result of the Vietnam draft lottery, one of them is inducted into the Army and subsequently enjoys poor outcomes, but not the other. This is a special form of Dworkin’s notion of *brute luck*, which represents a situation where the individual cannot reasonably impact the probability of an event taking place. This kind of luck can occur at any time over a life course. Vallentyne (2002) distinguishes two types of brute luck. *Initial brute luck* is defined as the set of factors that influence lifetime prospects up to the moment when individuals can be considered responsible for their choices and decisions. This roughly corresponds to Arneson (1990)’s idea of a “*canonical moment*” where individuals become responsible for their choices and preferences. By contrast, *later brute luck* denotes the luck factors that affect individual outcomes after the canonical moment. Our example illustrates *later brute luck*.

Fourth, consider two individuals who both have to choose among two lotteries. The outcome of the first lottery is certain. The outcome of the second is random. Assume that individuals make different choices and end up with different outcomes. The occurrence of outcomes partly escapes individual control, although by making different choices, one can influence the occurrence of outcomes. This corresponds to Dworkin’s notion of *option luck*. This notion implies that risk is taken deliberately, is calculated, isolated, anticipated and avoidable.⁹ We assume it is the case in our example and refer to it as *informed option luck*.

2.1.2 The requisites of equality of opportunity

Whether (and how), from an egalitarian perspective, these different varieties of luck ought to be compensated has been the subject of numerous papers. Their main (unconsensual) conclusion is that not all types of luck singled out in the previous paragraphs call for full compensation.

Almost all authors would agree that *social background luck* should be fully compensated, resorting to the ‘*starting gate position*’ argument : some deep inequalities of life prospects related to economic and social circumstances of birth cannot be justified by appeal to merit and desert (Rawls 1971).¹⁰ By full compensation, we mean that justice requires that outcomes be equal regardless of *social background luck*, other things being equal.

A similar argument applies to the effects of *genetic luck* on individual outcomes. However, given the constitutive nature of *genetic luck*, compensation of its impact may conflict with other ethical values. Hence, it has been claimed that *genetic luck* should not be compensated, owing

⁹Lippert-Rasmussen (2001) and Fleurbaey (2001) emphasize the strong informational requirements that underlie the notion of option luck : option luck presupposes that agents share similar subjective and objective probabilities of outcome occurrence.

¹⁰See Swift (2005) for a discussion of the legitimacy of parental influence on child’s outcomes.

to the libertarian principle of self-ownership which states that agents are entitled to the full benefit of their natural personal endowments (e.g. intelligence, beauty, strength) (Nozick 1977, p.225). For some authors, this requirement should receive priority over other principles.¹¹

From a moral point of view, compensation for all forms of luck has also been contested on efficiency grounds. The cost of this compensation can obviously be quite high. Such compensation requires considerable (and costly) information on individual situations as well as strong redistribution which may lead to large distortions, in the case of option luck. If these costs are large enough, compensating for all forms of luck may diminish the overall well-being. This has led some authors to formulate a restricted requirement of justice, which only calls for the compensation of *initial brute luck* and avoids part of the cost of redistribution. According to Vallentyne (2002), justice only requires that the *initial* value of lifetime prospects be equal across individuals, where the initial value is computed at the onset of adulthood.¹² This requires compensating for *initial brute luck*. Of course, to the extent that *later brute luck* is related to *initial brute luck*, compensation for the latter implies (at least partial) compensation for the former. However, equalizing the value of *initial* lifetime prospects does not erase all the impact of *later brute luck* on individual outcomes and individual can still end up, *ex post*, with different outcomes as a result of brute luck. It simply makes sure that *later brute luck* is *ex ante* even-handed, that is decorrelated from circumstances.

Lastly, three distinct views are held regarding the compensation for *informed option luck*. To the extent that the risky outcomes of option luck are avoidable and result from individual choice, some authors have claimed that inequalities resulting from option luck should not be compensated, owing to the principle of natural reward which states that the consequences of individual choice should be maintained. Dworkin supports that idea. A second view, expressed for instance in Fleurbaey (1995), recommends full compensation of the outcomes of option luck. Two distinct arguments are given in favor of this proposal. First, the fact that pure option luck is an extremely restrictive notion of luck that is both very rarely met in practice and very difficult to assess empirically. Second, and more importantly, the fact that not compensating for the effect of option luck implies that small errors of choices involve disproportionate, and thus unfair, penalties for some individuals. Lastly, some authors take an intermediate stance and argue in favor of partial compensation of the effects of option luck. Vallentyne (2002),

¹¹For instance, Vallentyne (1997) claims that "there are several independent moral demands, that they include both a demand for self-ownership and a demand for equality, and that a very strong form of self-ownership [...] constrains the demands of equality".

¹²According to Vallentyne, one advantage of this procedure is that the *ex ante* evaluation of life-time prospects takes into account the cost redistribution. This construct is in many ways similar to the one developed by Arneson (1989) who suggests that equality of opportunity should be defined by the equality of "preference satisfaction expectations".

states that equity authorizes taxation of the results of good option luck to partly compensate individuals who suffered bad option luck. Fleurbaey (2008) proposes to decompose the outcome of option luck into two components : on the one hand, the individual choice of a specific lottery, which belongs to responsibility factors; on the other hand, the intrinsic randomness of outcome attached to any lottery, which is akin to luck. According to Fleurbaey, the former component should not be compensated while the latter should to some extent be compensated. This view is close to the one of Le Grand (1991) who claims that individuals who make similar gambling decisions should enjoy equal well-being and therefore be fully insured and that individuals who make different decisions should only bear the consequences of their gambling over their expected well-being.

2.2 A model of equality of opportunity : circumstances, effort and luck

Our purpose is to build an economic model of equality of opportunity flexible enough to accommodate the diversity of positions held in ethical debates. The previous section reveals the lack of agreement regarding how random factors should be accounted for in the definition of equality of opportunity. However three main conclusions emerge from this analysis. First, the idea that the *social background lottery* should be included in the set of *circumstances* seems beyond dispute. Second, the impact of some random factors should be respected, as in the case of effort. Lastly, there may exist some random factors that one wants to decorrelate from circumstances.

It seems clear from the above discussion that this model should incorporate three types of factors : *circumstances* denoted by a vector c capture the deterministic or random factors that are not considered a legitimate source of inequality; *effort* summarized by a scalar e includes the determinants of outcome, either deterministic or random that are seen as a legitimate source of inequality. Hence our notion of effort is broader than the usual definition and does not solely hinge on the criterion of individual responsibility. *Luck* is captured by a scalar l and this concept, from now on, only comprises the random factors that are seen as a legitimate source of inequality as long as they affect individual outcomes in a neutral way, given circumstances and effort. In this paper, we take a neutral stance on the question of what factors should count as *circumstances*, *effort* or *luck*, which, in our view, pertains to moral or political debates.

Define y the individual outcome and $F()$ its cumulative distribution function, which is assumed to be continuous. A *type* defines the set of individuals with similar circumstances. A

variety denotes the set of individuals with similar circumstances and effort.

At this point, an important remark is in order. Among the four parameters, income, circumstances, effort and luck, we only have three “degrees of freedom”. For instance, once the distribution of outcome, circumstances and effort have been identified, we can retrieve luck. More precisely, in this setting, the overall impact of luck can be measured by the level of outcome that an individual reaches for a given level of effort and circumstances, for by definition, lucky individuals are the one who enjoy higher outcomes. The distribution of outcome conditional on circumstances and effort, $F(y|c, e)$, measures how luck affects the outcomes of individuals of a given variety. It gives the odds of all possible outcomes that may *ex ante*¹³ occur for an individual of this variety, as a result of the influence of luck. Alternatively, without loss of generality, luck may be summarized by a scalar index l . In this case, let $Y(c, e, l)$ denote the outcome function. Again, by the very definition of luck, this function must be strictly increasing in l . An example of such an index l may be defined by the “rank” where the individual sits in the distribution of outcome conditional on her variety : $l = F(y|c, e)$.¹⁴ So arbitrarily defined, l measures the relative degree of luck, within a given variety. By construction, l is identically distributed across varieties, which does not imply that a given degree of luck is associated with similar outcomes regardless of circumstances and effort.

2.2.1 A strong criterion of EOP

In this context, what does equality of opportunity require ? Since inequalities related to effort are morally acceptable, equality of opportunity should only apply among individuals with similar effort. Precisely, individuals with similar effort should face similar prospects for outcome, regardless of their circumstances. This is equivalent to say that given effort the distribution of outcome should not depend on circumstances. In turn, this requires that luck be even-handed, in the sense that the distribution of (the results of) luck is independent of circumstances. This criterion can be formalized by the following definition.

DEFINITION (EOP-S)

Equality of opportunity is satisfied iff: $\forall(c, c') \forall e, \quad F(\cdot|c, e) = F(\cdot|c', e)$.

One should emphasize that EOP-S only provides a formal definition of equality of opportunity. This formal definition is compatible with a variety of substantive conceptions of equality of opportunity, depending on one’s view of how all the concrete determinants of outcome should

¹³See below, § 3.4 the discussion about the *ex ante* and *ex post* approach.

¹⁴A similar definition is adopted in Fleurbaey (2008).

be classified as circumstances, effort and luck. As already mentioned, the criterion of individual responsibility, put forward by Cohen (1989), Arneson (1989) and Roemer (1993), offers a moral principle that can serve to define effort. This perspective is, however, in no way essential to our analysis. As suggested by the discussion in section 2.1, alternative principles, such as the principle of self-ownership, may serve to define our generic notion of effort. What matters to our analysis is that inequalities originating in differential effort are seen as legitimate and do not call for compensation. To give another illustration, if we consider that effort includes talent and ability, definition EOP-S leads to the Rawlsian conception of “*fair equality of opportunity*”, defined as a situation where “*those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system*” (Rawls 1971, p.63).

The present model also allows to analyze equality of opportunity in the presence of option luck. In this case, the effort variable should comprise the individual preference towards risk that condition her choice of a particular lottery. Under this interpretation, the condition in EOP-S implies that individuals with similar preferences should face similar lotteries independently of their circumstances. This is consistent with the view held by Dworkin. The compensation for option luck suggested by Le Grand (1991) would amount to place further restrictions on the effect of luck, conditional on effort, *i.e.* it would require that $F(\cdot|c, e)$ be a mass point distribution, at the expected value of well-being for a given choice of lottery.

Finally, two particular cases of this general model can be mentioned. First, a model where outcomes are only influenced by circumstances and luck as in the analysis of schooling outcomes undertaken in Jencks et al. (1972).¹⁵ In this case, EOP-S degenerates to the requirement that the distribution of outcome conditional on circumstances, $F(y|c)$ should be independent of c . Luck can lead to differences in individual outcomes as long as it remains neutral with respect to circumstances.

The second model corresponds to the world described by Roemer where luck is not considered *per se*. As the author admits (Roemer et al., 2003, p541), in his model, “*luck [...] unfortunately [...] appears as effort, because effort [is measured] as a residual after circumstances are accounted for*”. In that case, $F(\cdot|c, e)$ degenerates to a point-mass at $y(c, e)$. EOP-S requires that the income $y(c, e)$ be independent of circumstances.

¹⁵This model can be illustrated by the following citation. ‘*The main determinants of occupational and income success in society lie not with the amount of education that a person gets, but with what he brings to school with him, such as his genes, his “inherited” intelligence, his family background, and of course, his “luck” or lack of it*’ (Jencks, Smith, Acland, Bane, Cohen, Gintis, Heyns & Michelson 1972)

2.2.2 EOP-S and the conception of effort

The definition EOP-S is compatible with a variety of views on how effort should be defined and measured. In the literature, as well as in empirical applications, two main views have been taken.

A first possibility is the *absolutist* view of effort, which amounts to assume that effort can be defined in itself, without reference to circumstances. Under this view, the distribution of effort may differ across types.¹⁶ This view has been supported for instance by Barry¹⁷. It is not indisputable. As Roemer argues, if we are to take seriously the idea that individuals are not responsible for their circumstances, the definition of effort needs to be purged of any residual influence of circumstances. This leads to a *relativist* conception of effort. According to this view effort should *by construction* be distributed independently of circumstances : “*The choice of a degree of effort (as measured by the percentile of effort levels within a type) as the relevant metric for how hard a person tried, is justified by a view that, if we could somehow disembody individuals from their circumstances, then the distribution of the propensity to exert effort would be the same in every type*”.¹⁸ The choice between these two views is not an empirical matter but reflects different ethical conceptions of effort.¹⁹ Of course, not everyone would subscribe to the relativist view of effort. It is far from obvious that Roemer’s argument carries over to the case where effort also includes non-responsibility factors such as genetic or option luck.

Let e^A and e^R denote respectively absolute and relative effort. Roemer’s conception amounts to define e^R as a function of e^A and circumstances c , $e^R(e^A, c)$, such that the distribution of e^R is independent of c . A natural candidate is $e^R(e^A, c) = G(e^A|c)$, the rank in the distribution of absolute effort conditional on circumstances c .²⁰

One may wonder whether endorsing the relativist view of effort allows to simplify the above definition of equality of opportunity. Indeed, since by construction the distribution of effort is the same across circumstances, it may be tempting to aggregate the condition in EOP-S over effort levels. However, even under this view of effort, one can show that the condition in EOP-S does not degenerate to a simpler condition involving only circumstances, unless we impose additional restrictions. This is established by the following proposition.

¹⁶To simplify the analysis, we assume that the support of the distribution of effort is independent of circumstances.

¹⁷See Roemer (1998).

¹⁸Roemer (1998, p.15)

¹⁹Along similar lines, see the example developed in Cohen (1989, pp917-921).

²⁰A technical condition is required for $e^R(e^A, c)$ to be properly defined. The distribution of e^A conditional on c should not exhibit any mass point.

PROPOSITION 1

Under the assumption $e = e^R$, $\text{EOP-S} \implies \forall(c, c'), F(|c) = F(|c')$.

Proof : $F(y|c) = \int_{e^R} F(y|c, e^R)dG(e^R|c)$. By construction of e^R , for all c , $G(e^R|c) = G(e^R)$. Hence, if EOP-S is satisfied under the relativist view of effort, $F(y|c) = \int_e F(y|e^R)dG(e^R)$ and is independent of c .

Hence, under the relative view of effort, a necessary condition for EOP-S is that the outcome distributions conditional on circumstances alone be equal across types. It is important to stress, however, that this condition is not sufficient. Hence, even in the case of relative effort, it is not possible to collapse definition EOP-S into a condition on the distribution of outcome conditional on c alone. To establish that the reciprocal of proposition 2 is not true, consider the following counter-example with two types c and c' . Within each type, e^R (resp. l) takes two values (L, H) (resp. (g, b)) with equal probability. The outcome function $Y(c, e, l)$ is given by :

		type c			type c'			
		effort e^R			effort e^R			
			L	H		L	H	
luck l	g	1	3		luck l	g	1	2
	b	2	4			b	3	4

For each level of effort, the distribution of outcome varies according to circumstances, so EOP-S is not satisfied, although the aggregate distribution over effort levels are identical.

Only under additional restrictions is it possible to simplify the condition defining EOP-S. These restrictions correspond to the model developed by Roemer.²¹ This model first assumes that the relevant notion of effort for defining equality of opportunity is relative effort. Second, it assumes that random factors should be ascribed either to circumstances or to effort and in no case to luck. Hence, outcome can be expressed as a function $Y(c, e^R)$ that only depends on relative effort and circumstances and EOP-S requires that individuals with similar relative effort should receive equal outcomes regardless of circumstances. The third hypothesis is that the outcome function is strictly increasing in effort. In this case, although effort is not directly observable, relative effort e^R can be inferred from the observation of outcome and circumstances. For an individual with circumstances c , e^R can be assessed by the rank p where she sits in the conditional outcome distribution, $F(y|c)$. Let $Q(p|c)$ denote the quantile function associated to the distribution $F(y|c)$ and defined by $p = F(Q(p|c)|c)$. EOP-S requires that $\forall(c, c'), \forall p \in$

²¹For a complete discussion of the conditions of identification of equality of opportunity in Roemer's model, see O'Neill, Sweetman & Van De Gaer (2000).

$[0, 1]$, $Q(p|c) = Q(p|c')$). Hence we have the following proposition.

PROPOSITION 2

Under Roemer's assumptions, $\text{EOP-S} \Leftrightarrow \forall(c, c'), F(|c) = F(|c')$.

2.2.3 A welfarist foundation

Now we explore another possible justification for EOP-S, in a welfarist setting inspired by Arneson (1989) and Vallentyne (2002), and in which welfare should be equal across individuals, only to the extent that they exercise the same degree of responsibility.

Arneson (1989) and Vallentyne (2002) propose to use the *expected value* of future prospects as the relevant metric for evaluating opportunities. This is coherent with the idea that, from an *ex ante* perspective, given individual circumstances, luck, and consequently outcomes, may be seen as random processes. In this context, equality of opportunity can be defined by the equality of the expected value of future prospects across individuals. To perform this, one can use a specific Von Neumann-Morgenstern utility function u and compute the expected utility of the opportunities for outcome offered to a given type. In this case, equality of opportunity is defined by the following proposition :

DEFINITION

Equality of opportunity is satisfied iff : $\forall(c, c'), \int_y u(y)dF(y|c, e) = \int_y u(y)dF(y|c', e)$.

However, the question of what utility function to choose remains problematic. Ideally, one would like the characterization of equality of opportunity to hold for a sufficiently broad class of utility functions. In the case where there is a natural ordering for the outcome under consideration, as is the case for income, it is reasonable to focus on monotone increasing and concave utility functions. In this context, it is obvious that the expected value of future prospects attached to different circumstances c will be equal, for all possible increasing concave utility functions if and only if the income distributions for these circumstances are equal. Hence, we get a welfarist foundation to EOP-S.

2.2.4 Two weak criteria of EOP

Even without placing any further restriction on the distribution of income, the situation characterized by EOP-S appears as a situation of *strong* equality of opportunity. This condition is very stringent and may not easily be satisfied in practice. Consequently one may wonder whether all situations where EOP-S is violated should be considered equivalent from the point

of view of equality of opportunity.²²

Assume that EOP-S is not satisfied for two types c , and c' and a given effort level. Two situations can arise. First, for all relative degrees of luck, one type, say c , always gets higher outcome than the other ($\forall l, Y(c, e, l) \geq Y(c', e, l)$ and the inequality is strict for some levels of l). Second, one type gets higher outcomes for some degrees of luck while the other type gets higher outcomes for other degrees of luck (for instance, unlucky type- c do better than unlucky type- c' but lucky type- c do worse than lucky type- c'). Now consider the hypothetical situation of someone who would be given the option to choose between circumstances c and c' , without knowing her degree of luck. This is a typical case of choice under risk. In the first case, the outcome distribution associated with (c, e) stochastically dominates the one associated with (c', e) . There is a large agreement among specialists of decision theory (Starmer 2000) to say that in this case, consistent preferences under risk, should lead to choose c over c' .²³ In the second case, there is no such unanimous preference for c over c' . The first situation represents a clear case of inequality of opportunity while the second corresponds to a weak form of equality of opportunity, where no set of circumstances yields an unambiguous advantage over the other.²⁴

Avoidance of first-order stochastic dominance is however a very weak requirement to define equality of opportunity and one may object that this condition is not restrictive enough. For instance, it would consider that equality of opportunity prevails between c and c' in the case where all agents of type c do worse than those of type c' except for the one with the highest relative degree of luck. One may provide a more restrictive definition of weak equality of opportunity by resorting to the criterion of second order stochastic dominance (SSD). It is indeed commonly assumed that the Von Neumann-Morgenstern utility function exhibits risk-aversion, which corresponds to the case where $u(\cdot)$ is concave. It is well known²⁵ that the expected value derived from a distribution $F(y|c, e)$ will be greater than the one derived from $F(y|c', e)$ for all increasing concave utility functions if and only if $F(y|c, e)$ stochastically dominates $F(y|c', e)$ at the second order, which we will denote $F(\cdot|c, e) \succ_{SSD} F(\cdot|c', e)$.²⁶

²²Empirically, this question seems highly relevant. For instance Dardadoni et al. (2005) and O'Neil et al.(2000) both test a condition close to EOP-S and conclude that it is violated. However, they do not offer a formal ranking criterion for situations in which this condition is violated.

²³This consensus reaches well beyond the Expected Utility Theory. There is also empirical support for that view. In some experiments (Birnbaum & Navarette 1998) individual choices may not accord with the first-order stochastic dominance criteria. However, as argued by Starmer (2000) this may occur in situations where stochastic dominance is opaque to the agent.

²⁴Van de Gaer (1993) and Van de Gaer et al. (2001) already proposed the use of stochastic dominance tools to define equality of opportunity in a different framework.

²⁵The requirement that choices under risk be consistent with the principle of second-order stochastic dominance (SSD) stated below does not require that the Von Neumann - Morgenstern axioms be satisfied. Machina (1982) proved that this property is valid under more general conditions within the context of non-expected utility theories.

²⁶Formally, the definition of second-order stochastic dominance is the following :
 $F(\cdot|c, e)$ strictly stochastically dominates $F(y|c', e)$ at the second order ($F(\cdot|c, e) \succ_{SSD} F(\cdot|c', e)$) iff: $\forall x \in$

Using these notations, we get the following definition of weak equality of opportunity under risk aversion:

DEFINITION (EOP-W1)

Weak Equality of opportunity under risk aversion is satisfied iff:

$$\forall c \neq c' \forall e, \quad F(\cdot|c, e) \not\prec_{SSD} F(\cdot|c', e).$$

The conception of equality of opportunity embodied in definition EOP-W1 states that, whatever the value of effort, it is never possible to rank the opportunities offered by different types using second order stochastic dominance. Hence this notion of equality of opportunity rules out situations in which, for at least one effort level, one type would provide an unambiguous advantage (in the sense of stochastic dominance) over other types. A special case of such advantage, which, admittedly, represents a strong form of inequality of opportunities, is the situation where one type would provide an advantage over other types for all values of effort. An even weaker form of equality of opportunity than EOP-W1 is to rule out the latter possibility. This leads to the following definition :

DEFINITION (EOP-W2)

Very weak Equality of opportunity is satisfied iff:

$\nexists(c, c')$ such that :

$$\forall e, \quad F(\cdot|c, e) \succeq_{SSD} F(\cdot|c', e)$$

and

$$\exists e \text{ such that } F(\cdot|c, e) \succ_{SSD} F(\cdot|c', e).$$

If EOP-W2 is satisfied, two situations can occur : either there exists no effort level for which one type is unambiguously advantaged (which corresponds to EOP-W1), or, if there exists one effort level for which one type is advantaged, than the same type is disadvantaged for at least one other effort level. In both cases, no type can be unambiguously preferred when effort has not been chosen yet. Hence, although it is a weaker requirement, EOP-W2 still represents an appealing notion of equality of opportunity.

$$\mathbb{R}_+, \int_0^x F(y|c, e)dy \leq \int_0^x F(y|c', e)dy \text{ and } \exists x \mid \int_0^x F(y|c, e)dy < \int_0^x F(y|c', e)dy.$$

3 Equality of opportunity : Empirical identification

If circumstances and effort are observable, it is straightforward to examine whether the requirements of EOP-S, EOP-W1 or EOP-W2 are empirically satisfied. However, in many cases, some of that information may be missing. In this section we discuss to what extent equality of opportunity can be assessed when effort or circumstances are not fully observable to the empirical analyst and derive empirically tractable *implementation criteria* (IC) for equality of opportunity. We first consider the case where only outcome and circumstances are fully observable.

3.1 Unobservability of effort

We only observe the distribution of outcome conditional on circumstances $F(y|c)$ while in a complete information setting the conditions for equality of opportunity are stated in terms of the distribution of outcome conditional on circumstances and effort, $F(y|c, e)$. Letting $G()$ denote the cumulative distribution of e , the relationship between the two distributions is given by:

$$F(y|c) = \int_e F(y|c, e) dG(e|c). \quad (1)$$

Here, one may think of assessing equality of opportunity on the sole basis of observable conditional distributions. One obvious candidate for an implementation criterion of EOP is the equality of these conditional distributions.

DEFINITION (IC1)

Implementation criterion IC1 is satisfied iff : $\forall(c, c'), F(|c) = F(|c')$.

Unfortunately, without additional conditions, IC1 does not provide an implementation condition for EOP-S.

One interesting case arises when the distribution of effort is independent of c . This corresponds to two distinct situations. First, the analyst may adopt a relativist view of effort, in which case the independence is granted by construction. Second, the analyst may adopt an absolutist view of effort. In the latter situation, nothing guarantees the independence, but it still may be satisfied empirically. When adopting the relativist view, an interesting question, which has not been addressed in the normative literature, is whether effort should also be purged of the influence of luck. Two opposite arguments seem relevant from an ethical perspective. On the one hand, luck is beyond individual responsibility, which suggests that its influence on effort should be nullified. In this case, relative effort would be defined by $e^R = G(e | c, l)$. On the other hand, luck differs from circumstances in the sense that we want to compensate circum-

stances but not luck, as long as it remains neutral. If luckier individuals exercise more effort, their higher outcome may be seen as legitimate. In this case, effort should only be purged of the effect of circumstances. These two points of view lead to different conceptions of equality of opportunity. However it is important to emphasize that they yield similar testable restrictions, in the case where effort is not observable. Consequently, without loss of generality, we will adopt the second view and define relative effort as effort net of the influence of circumstances, *i.e.* $e^R(e, c) = G(e|c)$.

If besides circumstances, outcome is determined by effort and luck, then contrary to what happens in the Roemer model, the rank in the distribution of outcome conditional on c can no longer serve to retrieve effort, even if we assume that outcome is strictly increasing in effort. In our framework, the rank reflects the joint impact of luck and effort. For this reason, we only have necessary (but not sufficient) conditions for EOP-S, which is summarized by the following proposition.

PROPOSITION 3

If $\forall c, G(e|c) = G(e)$ then EOP-S \implies IC1.

This result extends proposition 2. It emphasizes that IC1 is a necessary but not sufficient condition for EOP-S as long as effort is independent of circumstances.

We now turn to the assessment of weak equality of opportunity, as defined by EOP-W1. Consistent with the relativist view of effort, we assume that effort is independent of circumstances. The consequences of the unobservability of effort are more serious for asserting EOP-W1 than for EOP-S. In a nutshell, weak equality is defined by some inequalities, conditional on effort and circumstances, that do not survive well to integration over effort levels. Hence, in general, it is not possible to assess EOP-W1 without observing effort.

However, when effort is independent of circumstances, we get a sufficient condition for preserving EOP-W2, given by the following implementation criterion.

DEFINITION (IC2)

Implementation criterion IC2 is satisfied iff : $\forall c \neq c', F(\cdot|c) \not\prec_{SSD} F(\cdot|c')$.

This condition rules out situations of strict stochastic dominance across all possible pairs of types. It is sufficient to ensure the weakest conception of EOP.

PROPOSITION 4

If $\forall c, G(e|c) = G(e)$ then : IC2 is a sufficient condition for EOP-W2.

Proof : Assume that EOP-W2 is violated. We want to show that in this case, IC2 is not satisfied. First note that $\int_0^x F(u | c)du = \int_0^x \int_E F(u | c, e)dG(e|c)du$. Using Fubini's theorem and the assumption of independence of the distribution of e w.r.t c , we have $\int_0^x F(u | c)du = \int_E \int_0^x F(u | c, e)dudG(e)$. Violation of EOP-W2 implies that there exists (c, c') such that, for all e , we have : $\forall x \in \mathbb{R}_+$, $\int_0^x F(u | c, e)du \leq \int_0^x F(u | c', e)du$. And there exists at least one value of e such that this inequality is strict for some x . By integration, we have $\forall x \in \mathbb{R}_+$, $\int_0^x F(u | c)du < \int_0^x F(u | c')du$, and the inequality is strict for at least some x . Hence IC2 is violated.

3.2 Partial observability of circumstances

We now consider the case where effort and some circumstances are not observable. The vector of observable circumstances is denoted by c_1 and the vector of unobservable circumstances by c_2 . In this case, it is still possible to provide a necessary condition for EOP-S when effort is independent of circumstances without any other assumption. This condition is given by the criterion IC1, defined on observable circumstances. This is summarized by the following proposition.

PROPOSITION 5

If $\forall c, G(e|c) = G(e)$ then : EOP-S $\implies \forall(c_1, c'_1), F(\cdot|c_1) = F(\cdot|c'_1)$.

Proof : If EOP-S is satisfied, we have : $\forall(c_1, c'_1), \forall c_2, \forall e, \forall y, F(y|c_1, c_2, e) = F(y|c'_1, c_2, e)$.

Furthermore, if e is distributed independently of circumstances, integrating over values of e implies : $\forall(c_1, c'_1), \forall c_2, F(\cdot|c_1, c_2) = F(\cdot|c'_1, c_2)$

Integrating over values of c_2 implies : $\forall(c_1, c'_1), F(\cdot|c_1) = F(\cdot|c'_1)$.

Roemer has suggested a similar result in his own setting without luck. It is important to stress that our result does not require the independence of unobserved circumstances vis à vis observed circumstances is needed. In this setting, as in Bourguignon, Ferreira & Menendez (2007), observed circumstances will also capture the influence of any correlated unobserved circumstances.

One important implication of the above proposition is that when the set of circumstances is multi-dimensional, it may be possible to test for equality of opportunity by testing condition IC1 for each circumstance separately and ignoring the other circumstances. This may be particularly helpful in the presence of a small sample for which the partitioning based on the full set of circumstances would lead to very few observations for each type.

The reciprocal of 5 is however not true. Loosely speaking, proposition 5 simply states that the values of $F(\cdot|c_1, c_2)$ and $F(\cdot|c'_1, c_2)$ are equal on average, where the average is computed over values of c_2 . Of course, this does not imply that the CDF are equal for all values of c_2 .

Again, assessing EOP-W1 is not achievable under partial observability of circumstances. However a special case of EOP-W2 can be assessed, which corresponds to the case where no visible type c_1 dominates a visible type c'_1 , for all possible efforts, whatever the value of their unobservable circumstances c_2 . That is

$$\nexists(c_1, c'_1)|\forall(c_2, c'_2), \forall e, F(\cdot|c_1, c_2, e) \succ_{SSD} F(\cdot|c'_1, c'_2, e) \quad (\text{EOP-W3})$$

It is still possible to provide a sufficient condition for the avoidance of EOP-W3, which is given in the following proposition.

PROPOSITION 6

If $\forall c, G(e|c) = G(e)$ then IC2 is a sufficient condition for EOP-W3.

Proof : Negating EOP-W3 and integrating the dominance relation over effort and unobservable circumstances implies that $F(\cdot|c_1) \succ_{SSD} F(\cdot|c'_1)$. Hence if effort is independent of circumstances, IC2 is violated.

3.3 Ex ante versus Ex post

All conceptions of equality of opportunity discussed so far amount to equalize the prospects for outcome that individuals face, given their effort level. These prospects obviously refer to the *ex ante* distribution of outcomes, before the individuals know their degree of luck. Of course this *ex ante* distribution is never observable empirically. At best, what we can recover is the *ex post* distribution of outcome among individuals of a given variety. The empirical characterization of equality of opportunity undertaken in the previous sections rests on the assumption that the *ex ante* prospects offered to individuals of a given variety can be identified with the *ex post* distribution of outcomes among individuals of that type. In other words, what is required is a law of large numbers implying that the frequency distribution in the cross section be identical to the probability distribution for the individual.²⁷ This reduction hypothesis raises some empirical issues and requires specific conditions that, although unavoidable, deserve to be explicitly stated.²⁸

The first problem is common in statistical inference. In some cases, the size of the sample

²⁷See Judd (1985).

²⁸We thank an anonymous reviewer for drawing our attention to this problem.

may be too small to prevent a noticeable difference between the probability and the frequency. In testing stochastic dominance, we have to keep an eye on the size of the sample.

The second problem is more specific to the reduction hypothesis implicitly made thus far. By construction, all individuals within a variety are offered similar *ex ante* prospects. But this will only correspond to the *ex post* distribution of outcome if individuals independently draw from the *ex ante* distribution. In some cases, however there are reasons to expect outcomes to be correlated within a type. For instance, the children of farmers are more likely to become farmers themselves and thus, are exposed, in any given year, to common weather shocks. Hence, the *ex post* distribution of yearly income may differ from the *ex ante* distribution of prospects. In this simple example, one may think that using repeated observations of yearly income may help solve the problem. But in some cases, there may be some path-dependency at the type level. For instance, it can occur that a generation of farmers has been hurt by bad weather conditions at the beginning of their professional life, which forced them to sell their land to banks and to become tenants for their remaining life. As a result, the income level of the whole type has been lowered. In this case, although *ex ante*, their life chances of income may have been the same as for people living in cities (assumption), it is no longer true *ex post*. But in this case, one may wonder how relevant the *ex ante* distribution turns out to be.

In summary, in any empirical implementation, whatever the cautions one can take, we have to admit that it is an *ex post* view that it is tested. Is it annoying? Since *ex ante* EOP may be always true if we go sufficiently back in time, a robustly observed *ex post* view may defended as the right one as well. As Marc Fleurbaey (2008) wrote “*taking the ex-post point of view does not imply ignoring individuals’ ex ante life and sense of agency, it simply consists in looking at their ex ante life with the benefit of hindsight about the ex-post distribution of luck*”. This is exactly what we attempt to do in practical terms.

3.4 Summary

To summarize the main conclusions of this section, equality of opportunity can be empirically assessed using the conditions IC1 and IC2, in the case where effort or some circumstances are not observable. We need to compare the cumulative distributions of income, conditional on observed circumstances. Since IC1 is only a necessary condition for equality of opportunity in the general model, we can only draw firm conclusions in the case where the cumulative distributions are not found equal. This case indicates that equality of opportunity, as defined by EOP-S is violated. The situation where the cumulative distributions are found equal is only

indicative of equality of opportunity: we can only conclude to equality of opportunity if we are willing to consider that the determinants of outcome excluded from the circumstances only reflect either luck or effort.

Since it is also well-known that SSD is equivalent to generalized Lorenz dominance, IC2 requires to compare the generalized Lorenz curves associated with observed circumstances. When comparing two generalized Lorenz curves, three situations can occur: (a) the two curves are identical, (b) the two curves intersect, (c) one curve lies above the other. Case (a) is equivalent to the equality of the cumulative distributions, which has already been discussed. Case (b) entails very weak inequality of opportunity, as defined by EOP-W2 or W3. It also implies weak equality of opportunity in the sense of EOP-W1 if we are willing to assume that the determinants of outcome excluded from the circumstances resort to luck or to effort alone. Lastly, case (c) is suggestive of a deviation from equality of opportunity for two reasons : first, EOP-S is not satisfied; second we cannot even rule out the situation of strict dominance between two observed types. Case (c) corresponds to the situation of second-order stochastic dominance of one observable conditional distribution over the other. A special case of this situation is the case of first-order stochastic dominance. This case is worth investigating in its own right, since FSD signals a strong deviation from equality of opportunity with a full ranking of the opportunity sets offered to individuals with different circumstances whatever the attitude of the decision maker towards risk.

4 Empirical application : income in France, 1979-2000

In this section, we analyze equality of opportunities for income in France. To this end, we examine whether income distributions conditional on social origin are equal or exhibit stochastic dominance patterns. We first present the data and the statistical procedure used in the analysis. We then discuss the results.

4.1 Data

The data come from the French household survey "Budget des Familles" (BdF) conducted by the French national statistical agency (INSEE). Five waves of the survey have been collected (1979, 1984, 1989, 1994 and 2000), each on a sample of about 12,000 households. We use all available waves.²⁹ For each household, the data provide detailed information on all sources

²⁹We use sample weight to ensure sample-representativeness.

of income and expenditures and enable to identify the household's social background. Sample summary statistics are presented in appendix table A-1.

4.1.1 Main variables

All waves of the BdF data contain information on the social origin of heads of household and their spouse. Both are asked to report the one-digit occupational group of their two parents.³⁰ From these four variables, it would be possible to build a detailed classification of social origin. Given the size of our samples, the use of a detailed classification would lead to small sub-samples, and to inaccurate estimations of the income distributions conditional on social origin. For this reason, we only use information on the occupational group of the household head's to define individual circumstances. This leads to distinguish the following six social origins: farmers, artisans, higher-grade professionals, lower-grade professionals, non-manual workers and manual workers.³¹

The outcome variable we focus on is household standard of living. The BdF data provide a detailed record of all income sources including wage and labor income, asset income, transfers (pensions, unemployment benefits, child support, welfare benefits) as well as income and property taxes. We consider two measures of family income. The first one is primary income, which includes labor and asset income, and unemployment and pension benefits. The second one corresponds to disposable income and is equal to primary income plus redistributive transfers minus taxes. In both cases, we normalize family income by family size using the OECD equivalence scale. To make income measures comparable over time, income is expressed in constant terms (2002 Euros) using the consumer price index.

Consequently, our data set provide a comprehensive measure of one fundamental individual outcome : living standard. It also offer a characterization of an important determinant of this outcome, that most authors would agree to include among the set of relevant individual circumstances. On the contrary, it offers no measure of individual effort or luck and provides only an incomplete description of circumstances. Therefore, the empirical assessment of equality of opportunity that our data allow corresponds to the situation analyzed in section 3.2.

³⁰For every survey, except 1979, it is the occupational group when the respondent was 16. In 1979, it is the last occupational group of the parents.

³¹These six groups correspond to the French INSEE job classification. Children of artisans also include the children of small proprietors. The occupational groups of the 1979 survey have been recoded to account for the change in the occupational classification that occurred in 1982.

4.1.2 Sample selection rules

Within a given survey wave, changes in the social structure over time imply that the age composition will differ across groups of different social origin. For instance, the rise in the share of higher-grade professionals and the fall in the share of farmers implies that children of higher-grade professionals (respectively farmers) will be younger (resp. older) than the average. To avoid this composition effect, our sample is limited to households whose head was between 30 and 50 years old at the time of the survey. We also exclude households whose head was retired or student at the time of the survey. Another advantage of this sample selection rule is that household income will be more representative of their lifetime income (Grawe 2006). This leads to samples of about 4 000 households in each wave.

The early waves of the survey exhibit a high rate of non-response to the questions pertaining to income earned and taxes paid. In some case, information is missing for one or several income items. In others, respondents only report some income items in bracketed form. In these data, non-response cannot be considered random. It is correlated with the occupation of the head of household as well as with other socio-demographic characteristics, and is stronger for self-employed workers (farmers and artisans) than for wage earners. Hence, ignoring missing data would lead to a biased view of the income prospects conditional on social origin. For this reason, in case of non-response, household income has been imputed, using the simulated residuals method. For observations with missing or bracketed data, we predict income using an estimated income equation. This equation is estimated on those households who report an income (in level or in brackets), and income is regressed on observable characteristics (age, sex, occupational group, last diploma, consumption, nationality, family composition, geographic area of living ...)³². In case of missing data, we also draw a residual term that is added to the predicted income. This procedure is implemented to impute primary and disposable income. The income distribution estimated using this imputation procedure appear consistent with the distributions obtained from administrative data. This is true, in particular, for farmers' income that was especially badly reported in the first waves of the survey.³³

³²Detailed equations are available upon request.

³³As we will discuss, part of the results reported below are driven by changes in the income distribution of farmers, estimated from our data. Consequently, we were concerned that part of these evolutions may reflect spurious changes in income distribution related to changes in reporting behavior. Hence we compared the estimated income distribution of farmers to the ones obtained from agricultural national accounts. As documented in Lefranc *et al.* (2004), it turns out that our estimates are strongly consistent with those obtained from administrative sources.

4.2 Statistical Inference : General principles

Here, we explain the general principles of our statistical methodology. The details of the stochastic dominance tests we implement are presented in the Appendix. Our samples allow to build income distribution conditional on social origin and test whether conditions IC1 or IC2 are satisfied. Assessing equality or stochastic dominance relationships is a demanding exercise. It requires that the *entire* outcome distributions (or some integral of them) be compared for all possible circumstances. One should also bear in mind that these distributions need to be estimated and compared, in our case using samples of relatively small size. Hence, special attention must be paid to the statistical robustness of the conclusions drawn from sample data. In this context, performing parametric tests of stochastic dominance is likely to yield fragile conclusions. On the contrary, our empirical analysis rests on non-parametric stochastic dominance tests developed in Davidson and Duclos (2000), which lead to robust conclusions.

The empirical procedure we implement is the following. For all possible pairs of circumstances c and c' , we perform three tests independently : (1) we test the null hypothesis of equality of the distributions of types c and c' ; (2) we test the null of first-order stochastic dominance of the distribution of type c over type c' and *vice-versa*; (3) we test the null of second-order stochastic dominance of the distribution of type c over type c' and *vice-versa*.

Test (1) corresponds to condition IC1 and tests (2) and (3) to IC2. For any pair of types, we interpret the joint results of these tests in the way summarized below. Of course, this interpretation is only tentative and one should keep in mind the caveats discussed in section 3.4.

- If we fail to reject the null of test (1), we say that equality of opportunity is supported, since IC1 is satisfied.
- Else, if test (2) or (3) accepts dominance of one distribution over the other but not the other way round (*e.g.* $F(|c) \succeq_{SSD} F(|c')$ and $F(|c') \not\succeq_{SSD} F(|c)$) we say that equality of opportunity is violated, since neither IC1 nor IC2 are satisfied.
- Else, if test (3) rejects dominance of each distribution over the other (*(i.e.* $F(|c) \not\succeq_{SSD} F(|c')$ and $F(|c') \not\succeq_{SSD} F(|c)$)) we say that weak equality of opportunity is supported, since IC2 is satisfied but not IC1.
- Else, if test (2) or (3) conclude that the two distributions dominate each other (*(i.e.* $F(|c) \succeq_{SSD} F(|c')$ and $F(|c') \succeq_{SSD} F(|c)$)), we give priority to the result of test (1) since

it is a more powerful test of equality of distributions for any significance level. Hence, we say that only weak equality of opportunity is supported, since IC2 is satisfied but not IC1.

Lastly, one should note that, given our interpretation, conclusions of test (2) and (3) cannot contradict since the null of (2) is included in the null of (3). Thus the conjunction of the results of the three tests interpreted in this way cannot be inconsistent.

4.3 Results

In this section, we first report the results of the tests of equality and stochastic dominance, for the income distributions conditional on our partition of social origin. While we use all available waves, we only report in the main tables and figures the results for 1979 and 2000, since our discussion mostly focus on the initial and terminal waves. Results for other years are reported in the appendix.

Three main conclusions emerge from this analysis. First, equality of opportunity is not satisfied : for most pairwise comparisons of types, we find evidence of stochastic dominance relationships; overall, a clear hierarchy of the different groups of social origin emerges. Second, the *pattern* of inequality of opportunity is stable over time : the *relative* ranking of types remains almost constant across the period 1979-2000. Third, the *degree* of inequality of opportunity decreases over time : while the ranking of types is unchanged, the income distributions of the different types come closer together over the period.

We then examine what factors account for inequality of opportunity. As discussed in decision theory, two factors may contribute to stochastic dominance between two income lotteries : differences in the expected return of the two lotteries and differences in the degree of risk of the two lotteries. In our case, the expected return corresponds to the mean income for each type and risk corresponds to within type inequality. Our results indicate that the degree of risk is very similar for all types. This is true over the entire period. On the contrary, the returns differ markedly across types. Since the evolution of these returns is a key determinant of the narrowing of the income prospect gap, we finally analyze the determinants of the changes in these returns over time.

4.3.1 A reduction in the degree of inequality of opportunity

The cumulative distribution functions and the generalized Lorenz curves conditional on social origin are given in Figures 2 and 3 for 1979 and 2000. For 1979, a particularly clear ranking of social types emerges. Children of higher-grade professionals stand out as the most advantaged

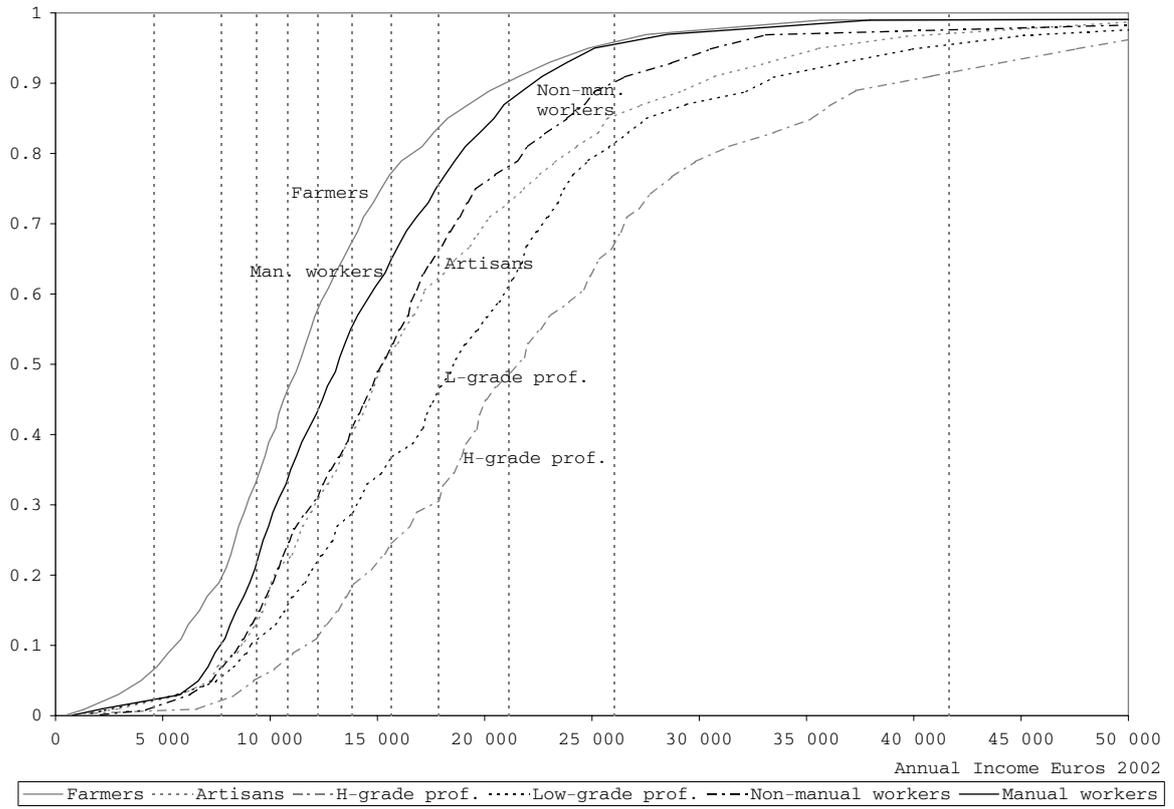
type : their conditional distribution dominates by far those of other social groups. Children of lower-grade professionals come next, followed by children of artisans and children of non-manual workers. In fact, the income distributions of the latter two groups seem very close, especially in the first half of the distribution. Lastly, at the bottom of the social hierarchy, come the children of manual workers and the children of farmers. The income distribution of the children of farmers, in 1979, is, by far, dominated by all other social backgrounds.

This “visual” ranking is strongly supported by the results of the tests of equality and stochastic dominance. These results are presented in Table 1, for primary and disposable income. In 1979, in all but one pair-wise comparisons, the equality of the conditional income distributions is rejected. Without ambiguity, this indicates that EOP-S is not satisfied. The only two types who apparently face equal opportunities are the children of non-manual employees and of artisans, although one should keep in mind, here and in the rest of the paper, that IC1 is only a necessary condition for EOP-S. Furthermore in all other pair-wise comparisons, the tests indicate that one distribution dominates the other. Hence we cannot even dismiss the case where EOP-W2 is violated. One should also note that in all these cases, stochastic dominance is satisfied at the first order, which implies that a ranking of social types can be achieved without assuming risk aversion. Lastly, in 1979, the impact of taxes and transfer on inequality of opportunity, is very limited. The gap between the generalized Lorenz curves is slightly lower for disposable income than for primary income (see figure 3), but stochastic dominance relationships are not affected by redistribution.

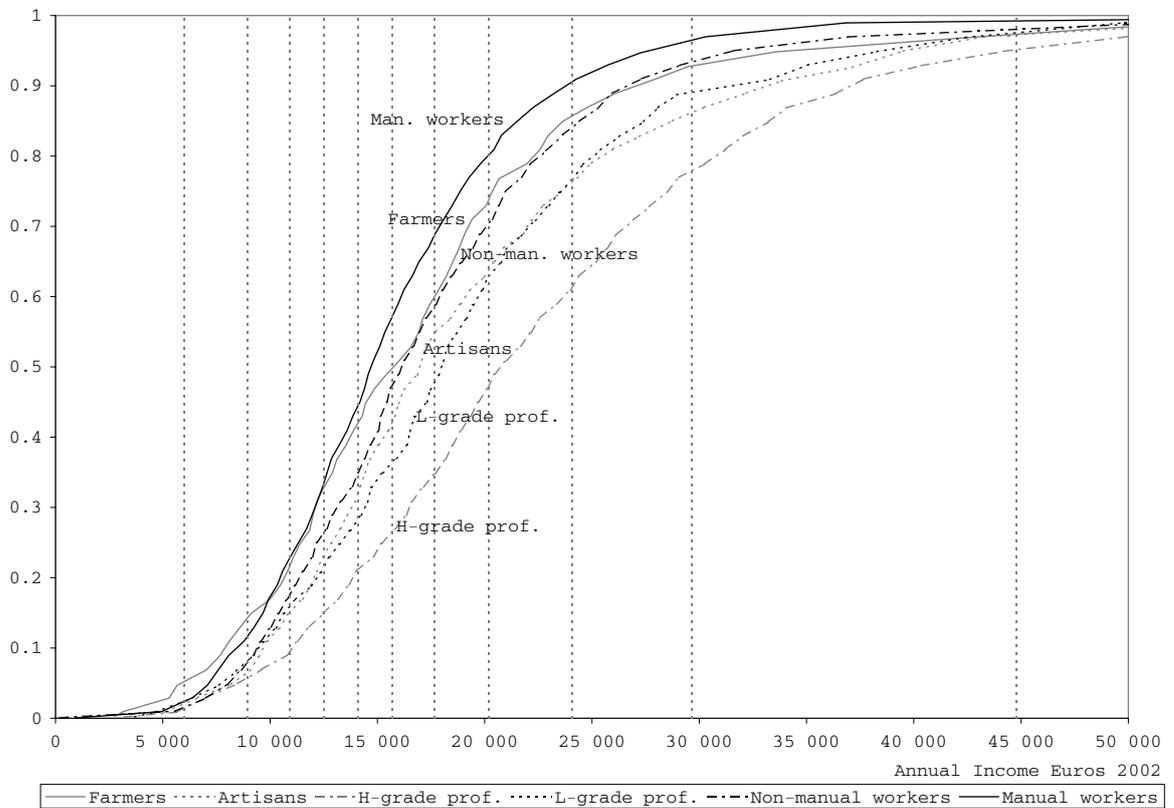
The pattern of stochastic dominance relationships exhibits small changes between 1979 and 2000. The results of the tests for primary income are given in Table 1. Four important features can be underlined. First, the dominant position of the children of higher-grade professionals remains unchallenged during the entire period: in every wave their income distribution dominates those of all other groups. Second, the hierarchy of intermediate groups tends to weaken. This is in great part due to an improvement of the relative ranking of the children of artisans : in 1979, this type was dominated by children of lower-grade professionals and their opportunities were equal to those offered to children of non-manual workers; this is no longer the case after 1994. In 2000, the opportunities offered to children of artisans are equal to those offered to children of lower-grade professionals and dominate those of children of non-manual workers. In this intermediate group, one can also notice a change in the relative ranking of children of lower-grade professionals and non-manual workers. As the table in appendix demonstrates in 1989 and 1994, the income distribution of these two groups are equal, although it is no longer

Figure 2: Income distributions by social background - disposable income

A- 1979

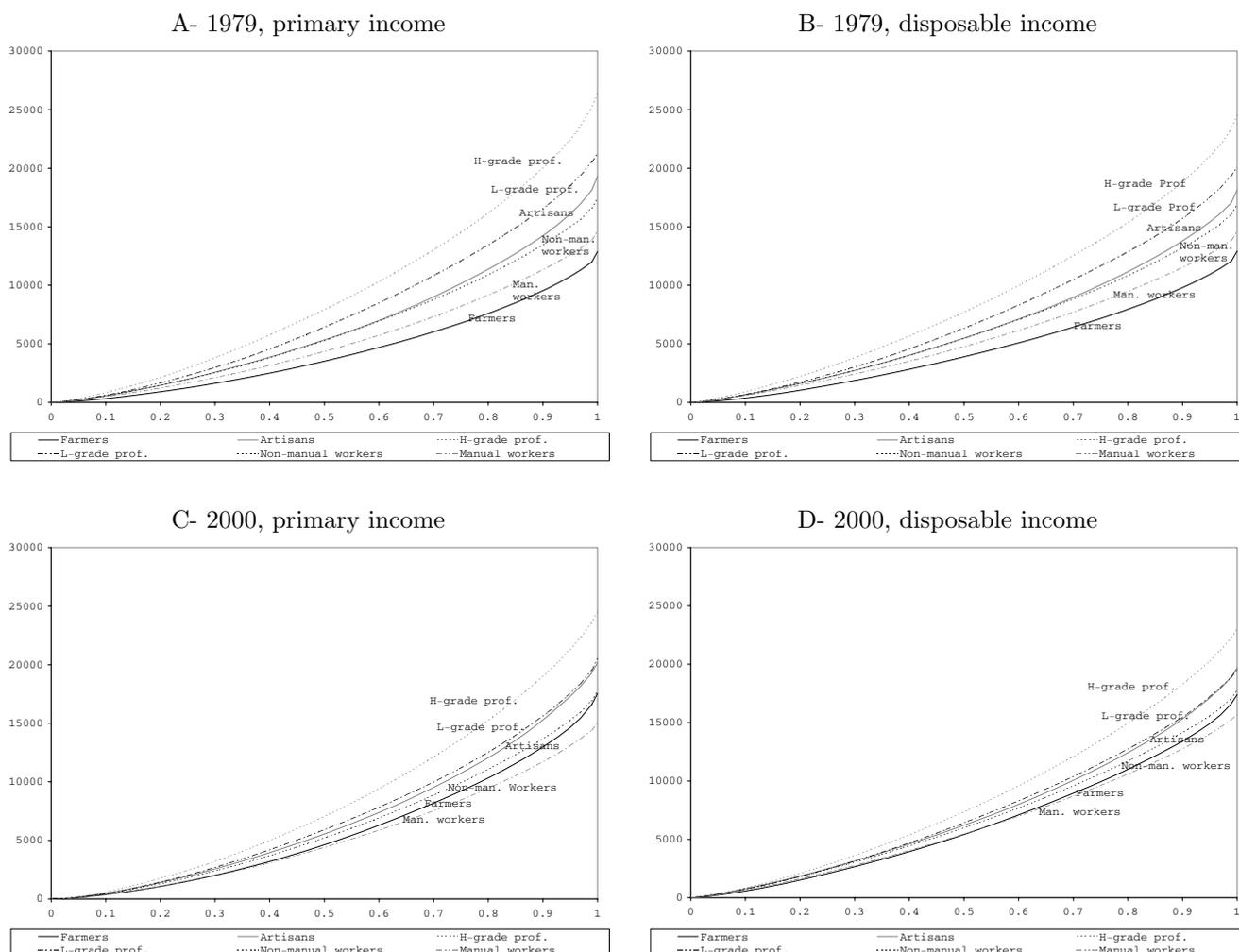


B- 2000



Notes : The occupational group refers to social origin. H-grade prof. : higher-grade professionals; L-grade prof. : lower-grade professionals; Non-man. workers : non-manual workers.

Figure 3: Generalized Lorenz curves in 1979 and in 2000



Notes : Income in Euros 2002 is represented on the *y*-axis. The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

true in 2000. The third important phenomenon that takes place over this period occurs at the bottom of the hierarchy. In 2000 the conditional distribution of the children of farmers dominates the distribution of children of manual workers. This group is now dominated by all the others social backgrounds. Overall, a three-levels hierarchy persists over the entire period. It is dominated by the children of higher grade professionals. In the middle comes an intermediate group that includes the children of lower-grade professionals, artisans and non-manual workers. At the bottom, come the children of farmers and manual workers.

While the ranking of social backgrounds has changed, one may nevertheless be tempted to conclude that equality of opportunity has not made any good progress over the period 1979-2000. In all waves, equality of opportunity is rejected in at least 80% of all pair-wise comparisons. In

Table 1: Stochastic dominance tests

A- Primary Income						
1979	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	=	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁
2000	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	?	> ₁
Artisans	-	-	< ₁	=	> ₁	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁
B- Disposable Income						
1979	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	=	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁
2000	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	> ₁
Artisans	-	-	< ₁	=	> ₁	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁

Notes : The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

Each element in the table indicates the result of the comparison of the income distribution of the groups in row and column using the tests presented in section 4.2 :

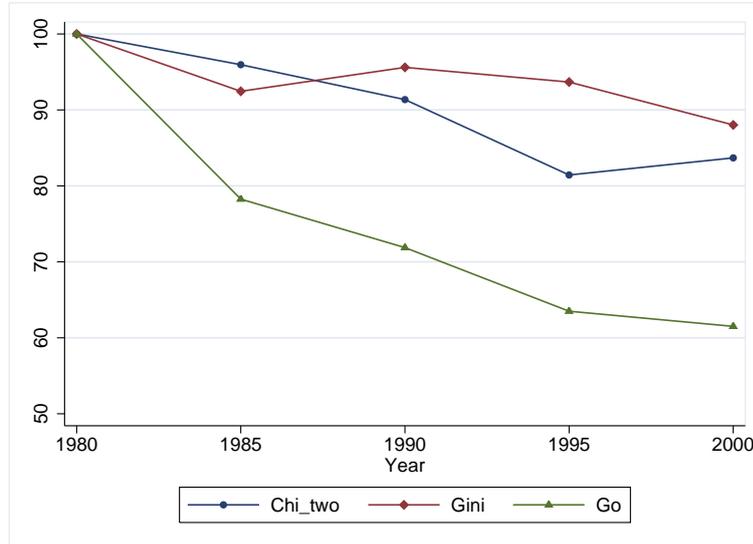
>_i: the row dominates the column for order *i* stochastic dominance ;

<_i: the column dominates the row for order *i* stochastic dominance;

=: the distributions are equal;

?: the distributions cannot be ranked using first and second order stochastic dominance.

Figure 4: Indexes of inequality of opportunity, income inequality and social immobility 1979-2000



Chi two: Index of social immobility, Gini: index of income inequality, GO: index of inequality of opportunity.

2000, IC1 is only satisfied in one case and IC2 prevails in another one. Hence, a strict ranking of all types is almost always possible.

IC1 is clearly violated in France. Given proposition 2, this implies that EOP-S is violated, for anyone who agrees that the fraction of any income determinants that is correlated with father's occupation should be treated as a circumstance and not as effort, in the practical definition of equality of opportunity.³⁴ IC2 is also clearly violated. Then under the same proviso, given proposition 3, we cannot even dismiss that one type is dominated by some other type for every level of effort.

Despite this stability in the ranking of social backgrounds, the comparison of the generalized Lorenz curves for the first and the last wave (see figure 3) clearly indicates that the income distributions of the different types have come closer together between 1979 and 2000. This suggests that the change at work is more cardinal than ordinal. While inequality of opportunity continues to prevail, the degree of inequality of opportunity seems to weaken.

A cardinal assessment of this decrease in inequality of opportunity is now proposed in Figure

³⁴This would apply to determinants as diverse as hours of work, work diligence, genes, preferences, knowledge, education, connections, beliefs,...

4. We compute an inequality of opportunity index consistent with our stochastic dominance approach and derived from the Gini coefficient. The detailed presentation of this so-called Gini-opportunity index (GO) and its properties is given in Lefranc et al. (2008).³⁵ Over the period we observe a sharp drop of the GO index at a decreasing rate. The level of inequality of opportunity in 2000 is 40% lower than in 1979. The same trend does not occur for the income inequality measured by the Gini coefficient which shows a much less drop. Interestingly, social immobility measured by the Chi-two index³⁶ witnesses a different evolution. After a sharp drop up to 1995, social immobility starts to rise again slightly in the last period contrary to the GO. This divergence demonstrates that changes in inequality of opportunity do not only mirror changes in social immobility. We expand on that point in the last subsection.

We now turn to the analysis of this cardinal change. To do so, we analyze the degree of risk and the return attached to the different conditional income distributions.

4.3.2 The risk of social lotteries

As already discussed, stochastic dominance relationships among income lotteries can arise because of differences in their expected return or in their degree of risk. In the present context, the expected return is defined as the mean income conditional on social origin and the risk corresponds to within-type inequalities. Note that in the general model of section 2 the income lotteries offered to individuals correspond to the income distribution conditional on their circumstances and effort. Return and risk should be computed from these distributions. In our case, we do not observe individual effort. Hence, within-type inequalities will reflect the joint influence of luck and effort.³⁷ This remark should be kept in mind throughout this section.

The degree of risk of the different social lotteries can be analyzed using the Lorenz curves of the distributions of income conditional on social origin. Consider two lotteries A and B . The lottery A is less risky than B if its Lorenz curve is always above the curve of B . In this case, lottery A is said to Lorenz-dominate lottery B . If the two Lorenz curves are identical then the

³⁵The definition of this index comes as follows. Let us rank all types according to twice the area under its conditional GL curve, starting from the smallest one. For type s , whose population share is p_s , this area is equal to

$$\mu_s(1 - G_s), \quad (2)$$

with μ denoting the mean and the G the Gini coefficient. The Gini-opportunity index is defined by :

$$GO(x) = \frac{1}{\mu} \sum_{i=1}^k \sum_{j>i} p_i p_j (\mu_j(1 - G_j) - \mu_i(1 - G_i)). \quad (3)$$

³⁶ $CHI2 = \sum_{j=1}^{j=6} \frac{(p_{jk} - p_j p_k)^2}{p_j p_k}$ where p_{jk} represents the proportion of individuals of social origin j and social destination k .

³⁷Effort can nevertheless be seen as a source of moral hazard and is therefore akin to a form of risk.

two social lotteries are equally risky. Consequently, to compare the degree of risk of the different social lotteries, we can resort to the same testing procedure as for stochastic dominance.

Table 2 displays the results of the tests of Lorenz dominance. The Lorenz curves of the different social lotteries are very similar. For primary income, tests conclude to the equality of the Lorenz curves, in 9 comparisons (out of 15) in 1979, 5 in 1984, 12 in 1989, 11 in 1994 and 10 in 2000. The tests conclude to dominance in only one case in 1979, 3 in 1984, 2 in 1994 and 0 in 1989 and 2000. Hence, the different social lotteries exhibit very similar degrees of risk.

A slightly different picture emerges from the analysis of the Lorenz curves of disposable income, for the end of the period. First, in 2000 (as well as in 1994), the Lorenz curve of the children of manual workers dominates that of all most other groups, with the exception of the children of non-manual workers and lower-grade professionals. In other words, the income distribution for children of manual workers exhibits the smallest degree of risk. Second, at the end of the period, the income distribution of children of farmers and, to a lesser extent, artisans, tend to exhibit more risk than those of children of wage earners.

Overall, these results suggest that social origin mostly influences the distribution of outcomes as a scale factor. On the other hand, the combined influence of effort and luck is such that the relative prospects are roughly similar across types. This implies that income conditional on social origin can be represented by the following multiplicative model:

$$y_{ic} = E(y | c)\epsilon_i \quad (4)$$

Where y_{ic} represents income of individual i with social origin s , $E(y | c)$ is the income mean conditional on c and ϵ_i a random term whose distribution is independent of social origin.

The strong similarity in the degree of risk attached to the different social backgrounds explains why in most cases, first-order stochastic dominance is a sufficient criterion for ranking conditional income distributions.³⁸ In fact, in this context, equality of opportunity can be assessed by relying solely on comparisons of mean conditional incomes. This situation is a priori quite rare.

³⁸The rare exceptions arise when comparing children of non-wage earners to children of wage earners. Due to the higher risk in the distribution for the former group at the end of the period, expected return and degree of risk sometimes point to opposite directions. For instance in 1994, as will be discussed in the next section, the mean disposable income is 5% larger for children of artisans than for children of lower-grade professionals. However, the stochastic dominance test turns out to be inconclusive due to the higher degree of risk of the lottery of children of artisans.

Table 2: Lorenz dominance tests

A- Primary Income						
1979	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	?	=	=	=	<
Artisans	-	-	=	?	?	?
H-grade prof.	-	-	-	=	=	?
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=
2000	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	=	?	?	?	<
Artisans	-	-	=	=	=	?
H-grade prof.	-	-	-	=	=	=
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=
B- Disposable Income						
1979	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	?	=	=	<	<
Artisans	-	-	=	=	=	=
H-grade prof.	-	-	-	=	=	?
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=
2000	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	=	=	?	<	<
Artisans	-	-	=	=	=	<
H-grade prof.	-	-	-	=	=	<
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=

Notes : The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

Each element in the table indicates the result of the comparison of the Lorenz curves of the groups in row and column using the criterion of Lorenz dominance :

$>_i$: the row dominates the column for Lorenz dominance ;

$<_i$: the column dominates the row for Lorenz dominance;

=: the Lorenz curves are equal;

?: the Lorenz curves cannot be ranked using Lorenz dominance.

4.3.3 The expected return of social lotteries

We now compare the expected returns of the different social lotteries. These expected returns can be summarized by the mean income conditional on social origin.

Trends in conditional mean income Mean incomes conditional on social origin are given in Table 3. This table confirms the evolution apparent in Figure 3 : mean conditional incomes tend to converge between 1979 and 2000. The ratio between the mean income of the most advantaged group (children of higher-grade professionals) and the least advantaged one (children of farmers in 1979 and of manual workers in 2000) falls from 1.89 to 1.46 for disposable income and from 2.04 to 1.63 for primary income. The advantage of children of higher-grade professionals falls by one half relative to the least advantaged group and diminishes relative to every other social background. More generally, the increase in the mean income of the different social groups is inversely related to their initial rank.

Table 3: Mean income conditional on social origin or destination

A- Mean income conditional on social origin				
	Primary Income		Disposable Income	
	1979	2000	1979	2000
Farmers	12 874	17 541	12 914	17 395
Artisans	19 295	20 174	18 137	19 691
H-grade prof.	26 375	24 543	24 490	23 033
L-grade prof.	21 225	20 511	20 055	19 534
Non-man. workers	17 379	17 720	16 884	17 747
Manual workers	14 612	15 008	14 592	15 709
Mean Income	16 503	18 313	16 070	18 180

B- Mean income conditional on social destination				
	Primary Income		Disposable Income	
	1979	2000	1979	2000
Farmers	9 367	17 858	9 614	17 449
Artisans	17 090	19 833	15 797	18 972
H-grade prof.	28 513	30 642	26 104	27 604
L-grade prof.	19 048	20 185	18 304	19 672
Non-man. workers	14 009	13 327	13 999	14 199
Manual workers	12 264	13 113	12 738	14 357
Mean Income	16 503	18 313	16 070	18 180

Notes : incomes in 2002 Euros. In panel A, the occupational group refers to social origin; in panel B, to social destination. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

The increase in mean income is particularly strong for children of non-wage earners: children

of farmers increase their mean income by 34.6%; children of artisans by 8.5%. As a result children of non-wage earners improve their relative position relative to all other social backgrounds. This contributes to the fall in the degree of inequality of opportunity since the initial ranking of both groups was relatively low : in 1979, children of artisans ranked third and children of farmers ranked last. The results for non-wage earners should be interpreted with caution due to the high rate of non-response discussed in section 4.1.2. However, the growth in non-wage earners mean income reported here is very close to what is observed using non-declarative sources such as national accounts. Moreover, most of this increase occurs during the 1990's, a decade for which we performed very few imputations. Hence, it is most unlikely that the erosion of higher-grade professionals' position and the improvement of non-wage earner's lot documented here arise from a statistical artifact. They reflect important changes in the extent of inequality of opportunity that need to be analyzed.

Decomposition This fall in the dispersion of mean incomes conditional on social origin may be related to fairly distinct phenomena. First it may originate from an increase in social mobility: if mobility increases, conditional mean incomes converge, as the distributions of social class destinations come closer together. Second, it may come from a reduction in the dispersion of mean incomes conditional on social origin and social destination. We refer to the first effect as the mobility effect, and to the second as the return effect.

The return effect can itself be decomposed into two effects, an absolute and a relative one. The absolute effect corresponds to a reduction in the spread of mean incomes conditional on social destination. The relative effect is detected by computing the ratio of the mean income conditional on social origin and social destination over the mean income conditional on social destination. Not all occupations classified under the same destination yield the same income. Coming from a high social origin may help occupy better paid jobs within each social destination. This phenomenon may lead to differences in mean income conditional on social origin and destination, once destination has been controlled. If these differences decrease over time, a relative return effect has been detected.

Formally, the mean income in year t , conditional on social origin j can be expressed as:

$$\bar{y}_j^t = \sum_k \alpha_{jk}^t r_{jk}^t \bar{y}_k^t \quad (5)$$

where k denotes social destination, α_{jk}^t is the probability of destination k conditional on origin j , $r_{jk}^t = \frac{\bar{y}_{jk}^t}{\bar{y}_k^t}$ where \bar{y}_{jk}^t denotes the mean conditional on social origin and destination and \bar{y}_k^t

Table 4: Occupational group transition matrices

A- 1979							
	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers	Total
Farmers	.225	.074	.047	.116	.14	.396	.216
Artisans	.009	.234	.192	.219	.147	.195	.139
H-grade prof.	.009	.033	.51	.261	.133	.05	.069
L-grade prof.	.000	.045	.287	.401	.138	.127	.073
Non-man. workers	.006	.057	.144	.295	.197	.298	.111
Manual workers	.005	.083	.064	.209	.164	.474	.389
Total	.053	.093	.134	.218	.156	.343	1.00

B- 2000							
	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers	Total
Farmers	.218	.049	.100	.159	.146	.325	.105
Artisans	.004	.118	.205	.251	.191	.229	.122
H-grade prof.	.005	.044	.415	.318	.119	.097	.150
L-grade prof.	.001	.055	.222	.349	.175	.186	.085
Non-man. workers	.007	.056	.137	.275	.232	.290	.184
Manual workers	.005	.047	.066	.192	.198	.490	.351
Total	.028	.058	.166	.243	.184	.318	1.00

Notes : the table gives the distribution of social destination (in column) conditional on social origin. The rows and columns labelled 'Total' give the column (resp. row) marginal distribution. Example: in 2000 sample, 10.5% of the population are children of farmers, 2.8% are farmers, and 4.9% of the children of farmers are artisans. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

the mean conditional on social destination. A change in α reflects a mobility effect, a change in \bar{y}_k^t , an absolute return effect, and a change in r_{jk}^t a relative return effect. We begin with a description of the direction of these effects over the period we study before moving to a decomposition analysis *à la* Oaxaca-Blinder.

The evolution of social mobility is summarized by the mobility matrices given in Table 4. The matrices indicate a rise in social mobility for several social groups. For example, children of higher-grade professionals see their probability of becoming higher-grade professionals fall from 51% in 1979 to 44% in 2000, while in the meantime the proportion of higher-grade professionals in the total population has increased. Children of farmers experience a large increase in upward mobility : while their probability of becoming farmer remains unchanged, the probability that they become higher-grade or lower-grade professionals increases between 1979 and 2000 and the probability that they become manual or non-manual workers decreases. Again, the mobility effect accounts for part of the change in mean incomes conditional on social origin.

As documented in Table 3, mean incomes conditional on social destination tend to partly converge between 1979 and 2000. Higher-grade professionals experience slower income growth

Table 5: Relative return effect

A- 1979					
	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	.95	.82	.93	.96	.95
Artisans	.94	1.08	1.05	.98	1.13
H-grade prof.	1.31	1.09	1.15	1.35	1.33
L-grade prof.	1.39	.95	1.08	1.10	1.08
Non-man. workers	.89	1.03	.96	.97	1.06
Manual workers	1.03	.85	.95	.95	.97
B- 2000					
Farmers	1.07	.94	1.03	.96	.96
Artisans	1.06	1.06	1.05	.97	1.01
H-grade prof.	1.21	1.05	1.05	1.04	1.05
L-grade prof.	1.05	.98	.97	1.00	1.11
Non-man. workers	.80	.94	.99	1.05	1.04
Manual workers	.93	.87	.94	.97	.97

Notes : the table gives the mean income conditional on social origin and social destination divided by the mean income of destination.

than other groups and non-wage earners experience faster growth. This indicates that the return effect narrows the dispersion of mean incomes conditional on social origin. However, it only accounts for part of the evolution. The observed reduction in mean income gaps is much stronger when conditioning on social origin than when conditioning on social destination. For instance the ratio between the mean disposable income of higher-grade professionals and manual workers only falls from 2.04 to 1.92. When conditioning on social origin instead of destination, it falls from 1.67 to 1.46.

To detect a relative return effect, we compute the ratio of the mean income conditional on social origin and destination to the mean income conditional on social destination³⁹. Table 5 shows that for almost every social destination, coming from a farmer or blue-collar background represents a handicap in 1979. It is no more true for descendants of farmers in 2000. On the opposite, originating from a high or low grade professional background represents a significant advantage which has decreased sharply at least for the former category over the period. On the whole, the relative return effect may be registered as a source of decrease in EOP.

The contribution of the mobility, the absolute and the relative return effects to the observed evolution can be identified using the Oaxaca-Blinder (1973) decomposition applied to equation (5). As is well-known, the decomposition is not unique. To assess the contribution of changes in α , r or y , there are several possible choices of a reference period for the factors held constant.

³⁹Farmers have been excluded from social destinations because of problems of statistical inference.

Table 6: Decomposition of the evolution of the mean income

	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
	Change in mean income (1979-2000)					
	4481	1 553	-1 457	- 521	863	1 118
Mobility effect	704	-28	-906	-989	-167	-179
(%)	15	-1	62	189	-193	-16
Relative return effect	968	-95	-2154	-966	-355	-91
(%)	21	-6	147	185	-41	-8
Absolute return effect	2808	1677	1603	1435	1385	1388
(%)	62	107	-110	-275	160	124

Notes : Change in mean income in Euros 2002. The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

Taking the arithmetic average on all six possible decompositions, the change in mean income conditional on social origin between t and t' can be written as:

$$\Delta \bar{y}_j = \frac{1}{6} (\Delta \alpha_j A + \Delta \mathbf{r}_j B + \Delta \bar{\mathbf{y}}_j C) \quad (6)$$

with

$$\begin{aligned} A &= 2\bar{\mathbf{r}}_j^t \bar{\mathbf{y}}_k^t + \bar{\mathbf{r}}_j^t \bar{\mathbf{y}}_k^{t'} + \bar{\mathbf{r}}_j^{t'} \bar{\mathbf{y}}_k^t + 2\bar{\mathbf{r}}_j^{t'} \bar{\mathbf{y}}_k^{t'}; \\ B &= 2\bar{\alpha}_j^t \bar{\mathbf{y}}_k^t + \bar{\alpha}_j^t \bar{\mathbf{y}}_k^{t'} + \bar{\alpha}_j^{t'} \bar{\mathbf{y}}_k^t + 2\bar{\alpha}_j^{t'} \bar{\mathbf{y}}_k^{t'}; \\ C &= 2\bar{\alpha}_j^t \mathbf{r}_j^t + \bar{\alpha}_j^t \bar{\mathbf{r}}_j^{t'} + \bar{\alpha}_j^{t'} \bar{\mathbf{r}}_j^t + 2\bar{\alpha}_j^{t'} \bar{\mathbf{r}}_j^{t'} \end{aligned}$$

where $\bar{\alpha}_j$, $\bar{\mathbf{r}}_j$ and $\bar{\mathbf{y}}_k$ denote the vectors $(\alpha_{j1} \cdots \alpha_{jK})$, $(r_{j1} \cdots r_{jK})$ and $(\bar{y}_1 \cdots \bar{y}_K)$. The terms A , B and C arise from the averaging procedure. The first term on the right hand-side of equation 6 is the mobility effect, the second is the relative return effect and the third one the absolute return effect.

The results of the decompositions for the mean disposal incomes are given in Table 6. The mobility and relative return effects contribute negatively to the change in the mean conditional income for every category except for the children of farmers. On the contrary and as could have been expected every social origin displays a positive absolute return effect, while its importance is broadly comparable across social groups, the absolute return effect is much more important for the children of farmers.

This pattern of results shows that mobility between categories as measured by the mobility

effect and mobility within categories, measured by the relative return effect, are positively correlated. Adding these two effects in a broad mobility concept, the following ranking of social groups appears: children of farmers experience a large positive increase in their mean income due to mobility. Next, the children of artisans display a slight decrease due to the same effect. Children of non-manual workers witness a larger negative drop in their mean income and then finally the children of lower-grade professionals and higher-grade professionals are much more negatively hurt by mobility within and between groups.

5 Conclusion

While different ethical positions can be defended, regarding how to substantively define equality of opportunity, our analysis indicates that any definition relies on a partition of the determinants of individual outcomes into three distinct groups : effort, which includes the determinants that are seen as a legitimate source of outcome differences; circumstances, which consist of the determinants that should not lead, other things equal, to differences in outcome; luck, which comprises the random determinants that are seen as a fair source of inequality provided that they are even-handed, with respect to circumstances. In this perspective, the generic model of equality of opportunity developed in this paper appears as a general model, that encompasses a variety of specific conceptions of equality of opportunity, depending on the precise empirical characterization of the above three sets of factors. This model also makes clear that, once these different sets have been delineated, there may still be several ways to define equality of opportunity, which correspond to the strong and weak criteria introduced here. These two criteria are expressed in terms of stochastic dominance because luck is akin to a form of risk in the realization of income that cannot be ignored.

Whatever the precise conception adopted, empirically assessing equality of opportunity turns out to be a data-demanding exercise. To many, equality of opportunity appears as a more desirable social objective than equality of outcome because it takes into account the determinants of observed outcomes. The obvious drawback of this conception is that making equality of opportunity judgments ideally requires that all the relevant determinants of outcome be observable. Of course, this requirement will rarely be met and assessing equality of opportunity will most likely take place under imperfect information on the relevant determinants. In this paper, we exhibit two testable conditions of equality of opportunity that can be used in this context. Because these two conditions are expressed in terms of stochastic dominance, our general model, incorporating luck, shapes our empirical strategy in a very distinctive manner, compared to the

rest of the literature.

The first criterion gives a necessary condition for the strong form of equality of opportunity. The other is a sufficient condition for the weak form. Only in very restrictive cases is it possible to exhibit a necessary and sufficient condition. Hence, in almost all cases, imperfect information implies that equality of opportunity cannot be fully assessed. This is probably the price to pay if we are willing to develop a rich enough view. It is also important to emphasize that these two conditions still allow, in our empirical application, to provide a valuable assessment of equality of opportunity in France. Even with limited information on the determinants of outcomes, these conditions indicate that equality of opportunity is clearly violated. This conclusion would therefore remain unchanged if we had access to a richer data set.

On the empirical side, this paper reveals that social inheritance is a deeply rooted source of inequality in France over the period 1979-2000. Differences in social origin translate into significant gaps of living conditions. Equality of opportunity in income acquisition does not prevail, neither for primary income nor for disposable income. However, the degree of inequality of opportunity tends to decrease. During this period, the average gap between the most advantaged social group, the children of higher-grade professionals, and the least advantaged one fall by one half.⁴⁰ The explanation of this evolution is addressed in a companion paper (Lefranc et al. (2006)). We show that this reduction in inequality of opportunity does not arise from a decrease in the degree of transmission of economic advantage from one generation to the next. Over time, children tend to face more equal opportunities because of a fall in inequality of outcomes among their parents.

Lastly, this paper has underlined an important phenomenon: the risk of social lotteries appears very similar across the different groups of social origin. As a first approximation, the influence of social origin on opportunities for income, in France, can be summarized by a scale factor : individual income is determined by the product of a random variable - distributed independently of social origin- and the mean income conditional on social origin. Whether a similar determination of income opportunities is also at work in other countries is a question that would be worth investigating. Future research should also analyze the theoretical explanation and consequences of this important stylized fact.

References

Alesina, A. & Angeletos, G.-M. (2005). Fairness and redistribution, *American Economic Review* **95**(4): 960–980.

⁴⁰This conclusion confirms those of Vallet (2004) who notices a slight increase in social mobility in France.

- Alesina, A. & Glaeser, E. (2004). *Fighting Poverty in the US and Europe: A World of Difference*, Oxford University Press, Oxford.
- Alesina, A. & La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities, *Journal of Public Economics* **89**(5-6): 897–931.
- Arneson, R. (1989). Equality and equal opportunity of welfare, *Philosophical Studies* **56**: 77–93.
- Arneson, R. (1990). Liberalism, distributive subjectivism, and equal opportunity for welfare, *Philosophy and Public Affairs* **19**(2): 158–194.
- Beach, C. & Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares, *Review of Economic Studies* **50**(4): 723–735.
- Ben-Porath, E., Gilboa, I. & Schmeidler, D. (1997). On the measurement of inequality under uncertainty, *Journal of Economic Theory* **75**(1): 194–204.
- Benabou, R. & Ok, E. A. (2001). Mobility as progressivity: Ranking income processes according to equality of opportunity, *NBER Working Papers 8431*, National Bureau of Economic Research.
- Betts, J. & Roemer, J. (2006). Equalizing opportunity for racial and socioeconomic groups in the United States through educational finance reform, in L. Woessmann & P. Peterson (eds), *Schools and the equal opportunity problem*, MIT Press.
- Birnbaum, M. & Navarette, J. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence, *Journal of Risk and Uncertainty* **17**: 49–78.
- Bishop, J. A., Formby, J. P. & Thistle, P. D. (1992). Convergence of the south and the non-south income distributions, 1969-1979, *American Economic Review* **82**(1): 262–272.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates, *Journal of Human Resources* **8**: 436–455.
- Bourguignon, F., Ferreira, F. H. G. & Menendez, M. (2007). Inequality of opportunity in Brasil, *Review of Income and Wealth* **53**(4): 585–618.
- Broome, J. (1991). *Weighing Goods. Equality, Uncertainty and Time.*, Blackwell, Oxford.
- Checchi, D., Ichino, A. & Rustichini, A. (1999). More equal but less mobile ? Education financing and intergenerational mobility in Italy and the US, *Journal of Public Economics* **74**: 351–393.
- Checchi, D. & Peragine, V. (2005). Regional disparities and inequality of opportunity: The case of Italy, *IZA Discussion Papers 1874*, Institute for the Study of Labor (IZA).
- Cohen, G. A. (1989). On the currency of egalitarian justice, *Ethics* **99**: 906–944.
- Corneo, G. & Gruner, H. P. (2002). Individual preferences for political redistribution, *Journal of Public Economics* **83**(1): 83–107.
- Dardanoni, V., Fields, G., Roemer, J. & Sanchez Puerta, M. (2005). How demanding should equality of opportunity be, and how much have we achieved?, in S. L. Morgan, D. Grusky & G. Fields (eds), *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, Stanford University Press, Stanford.
- Dardanoni, V. & Forcina, A. (1999). Inference for Lorenz curve orderings, *Econometrics Journal* **2**: 49–75.
- Davidson, R. & Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* **68**(6): 1435–1464.

- Diamond, P. A. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment, *Journal of Political Economy* **75**(5): 765–766.
- Dworkin, R. (1981a). What is equality. part 1: Equality of welfare, *Philosophy and Public Affairs* **10**: 185–246.
- Dworkin, R. (1981b). What is equality. part 2: Equality of resources, *Philosophy and Public Affairs* **10**: 283–345.
- Fleurbaey, M. (1995). Equal opportunity or equal social outcome, *Economics and Philosophy* **11**(1): 25–56.
- Fleurbaey, M. (2001). Egalitarian opportunities, *Law and Philosophy* **20**(5): 499–530.
- Fleurbaey, M. (2006). Assessing risky social situations, *mimeo* .
- Fleurbaey, M. (2008). *Fairness, Responsibility and Welfare*, Oxford University Press, Oxford.
- Fleurbaey, M. & Maniquet, F. (2007). Compensation and responsibility, in K. Arrow, A. Sen & K. Suzumura (eds), *Handbook of Social Choice and Welfare*, Vol. 2, North-Holland, Amsterdam, chapter 21.
- Foster, J. & Shorrocks, A. F. (1988). Poverty orderings and welfare dominance, *Social Choice and Welfare* **5**: 179–198.
- Gajdos, T. & Maurin, E. (2004). Unequal uncertainties and uncertain inequalities: an axiomatic approach, *Journal of Economic Theory* **116**(1): 93–118.
- Goux, D. & Maurin, E. (2003). On the evaluation of equality of opportunity for income: Axioms and evidence, *mimeo CREST* .
- Grawe, N. (2006). Lifecycle bias in estimates of intergenerational earnings persistence., *Labour Economics* **13**(5): 551–570.
- Hammond, P. J. (1981). Ex-ante and ex-post welfare optimality under uncertainty, *Economica* **48**(191): 235–250.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B. & Michelson, S. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*, Basic Books, New York.
- Judd, K. L. (1985). The law of large numbers with a continuum of iid random variables, *Journal of Economic Theory* **35**(1): 19–25.
- Knight, F. H. (1935). *The Ethics of Competition and Other Essays*, Harper & Bros, New York.
- Kodde, D. A. & Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions, *Econometrica* **54**(5): 1243–1248.
- Le Grand, J. (1991). *Equity and Choice*, Harper Collins, London.
- Lefranc, A., Pistolesi, N. & Trannoy, A. (2004). An equalization of opportunities for income acquisition in France?, *mimeo THEMA* .
- Lefranc, A., Pistolesi, N. & Trannoy, A. (2006). The link between inequality of opportunity for income acquisition and income inequality: the French example 1977-1993?, *mimeo THEMA* .
- Lefranc, A., Pistolesi, N. & Trannoy, A. (2008). Inequality of opportunities vs. inequality of outcomes: Are western societies all alike ?, *Review of Income and Wealth* pp. 513–543.
- Lippert-Rasmussen, K. (2001). Egalitarianism, option luck and responsibility, *Ethics* **111**(3): 548–579.

- Lippert-Rasmussen, K. (2005). Justice and bad luck, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
<http://plato.stanford.edu/archives/sum2005/entries/justice-bad-luck/>
- Machina, M. (1982). Expected utility analysis without the independence axiom, *Econometrica* **50**(2): 277–324.
- Marshall, G., Swift, A., Routh, D. & Burgoyne, C. (1999). What is and what ought to be: Popular beliefs about distributive justice in thirteen countries, *European Sociological Review* **15**(4): 349–367.
- Nozick, R. (1977). *Anarchy, State and Utopia*, Basic Books.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets, *International Economic Review* **14**: 693–709.
- O’Neill, D., Sweetman, O. & Van De Gaer, D. (2000). Equality of opportunity and kernel density estimation: An application to intergenerational mobility, *Advances in Econometrics : A research Annual* **14**: 259–274.
- Rawls, J. (1971). *A theory of justice*, Harvard University Press, Cambridge, MA.
- Roemer, J. (1993). A pragmatic theory of responsibility for the egalitarian planner, *Philosophy and Public Affairs* **22**(2): 146–166.
- Roemer, J. (1998). *Equality of Opportunity*, Harvard University Press, Cambridge.
- Roemer, J., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S. P., Lefranc, A., Marx, I., Page, M., Pommer, E., Ruiz-Castillo, J., San Segundo, M. J., Tranaes, T., Trannoy, A., Wagner, G. & Zubiri, I. (2003). To what extent do fiscal regimes equalize opportunities for income acquisition among citizens ?, *Journal of Public Economics* **87**: 539–565.
- Schuetz, G., Ursprung, H. & Woessman, L. (2005). Education policy and equality of opportunity, *CESifo working paper* (1518).
- Sen, A. (1985). *Commodities and Capabilities*, North-Holland, Amsterdam.
- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, *Journal of Economic Literature* **38**(2): 332–382.
- Swift, A. (2005). Justice, luck and the family : The intergenerational transmission of economic advantage from a normative perspective, in S. Bowles, H. Gintis & M. Osborne (eds), *Unequal Chances : Family Background and Economic Success*, Russel Sage Foundation and Princeton University Press, Princeton NJ, chapter 9, pp. 256–276.
- Vallentyne, P. (1997). Self-ownership and equality: Brute luck, gifts, universal dominance, and leximin, *Ethics* **107**(2): 321–343.
- Vallentyne, P. (2002). Brute luck, option luck and equality of initial opportunities, *Ethics* **112**: 529–557.
- Vallet, L.-A. (2004). Change in intergenerational class mobility in France from the 1970s to the 1990s and its explanation: An analysis following the CASMIN approach, in R. Breen (ed.), *Social Mobility in Europe*, Oxford University Press, Oxford, pp. 115–147.
- Van De Gaer, D. (1993). Equality of opportunity and investment in human capital, *Ph.D dissertation*, Katholieke Universiteit Leuven.
- Van De Gaer, D., Schokkaert, E. & Martinez, M. (2001). Three meanings of intergenerational mobility, *Economica* **68**(272): 519–538.
- Wolak, F. A. (1989). Testing inequality constraints in linear econometric models, *Journal of Econometrics* **41**: 205–235.

Appendix

A- Statistical tests

Stochastic dominance relationships can be easily expressed and statistically tested by making use of poverty indices. Formally the poverty index at the order g , for a poverty line z , is defined by :

$$D^g(z) = \int_0^z (z-x)^{g-1} dF(x)$$

For example $D^1(z)$ is the proportion of people whose income is below z ; $D^2(z)$ measures for the mean poverty gap, *i.e.* the mean amount that should be given to people below the poverty line z , to reach this threshold.

The link between poverty indices and stochastic dominance has been established by Foster and Shorrocks (1988). They have shown that stochastic dominance at the order g of distribution F_A over F_B is equivalent to the situation where the poverty index of order g for F_A is smaller than the poverty index for F_B , for all poverty lines. Letting \succeq_{SD_g} denote stochastic dominance at the order g , we have, for $g \in \mathbb{N}_+$:

$$F_A(x) \succeq_{SD_g} F_B(x) \Leftrightarrow \forall z \in \mathbb{R}_+ D_A^g(z) \leq D_B^g(z)$$

Consequently, tests of stochastic dominance are equivalent to test of inequality for poverty indices. The procedure adopted here follows Davidson and Duclos (2000) and consists in testing such inequalities for a fixed number, k , of poverty lines. The poverty lines used here are the deciles and 95th percentile of the overall household income distribution.⁴¹

For a given poverty line z , an unbiased and asymptotically normal estimator for the poverty index $\hat{D}^g(z)$ is given by⁴² :

$$\hat{D}^g(z) = \frac{1}{N(g-1)!} \sum_{i=1}^N (z-x_i)^{g-1} I(x_i \leq z)$$

where i denotes the i^{th} observation, N denotes the sample size and $I(\cdot)$ is an indicator function equal to 1 when its argument is true, 0 otherwise. For some fixed set of poverty lines $\{z_1, \dots, z_k\}$, let $\hat{\mathbf{D}}^g$ denote the vector of poverty indices $(\hat{D}^g(z_1) \dots \hat{D}^g(z_k))$ and Σ its asymptotic variance-covariance matrix⁴³.

The hypothesis of dominance at the order g between two distributions F_A and F_B can be expressed in vectorial notations. Let $\boldsymbol{\delta} = (\mathbf{D}_B^g - \mathbf{D}_A^g)$ the difference of the vectors of poverty indices. Stochastic dominance at the order g can be test as : $\mathbf{H}_0 : \boldsymbol{\delta} \in \mathbb{R}_+^k$ versus $\mathbf{H}_1 : \boldsymbol{\delta} \notin \mathbb{R}_+^k$. The test statistic is constructed from the estimated vector $\hat{\boldsymbol{\delta}} = (\hat{\mathbf{D}}_B^g - \hat{\mathbf{D}}_A^g)$, whose asymptotic variance-covariance matrix is given by $\frac{\Sigma_A}{N_A} + \frac{\Sigma_B}{N_B}$ under the hypothesis of independence between distributions A and B .

The null hypothesis is defined by a set of k constraints. Two approaches can be followed to conduct this test. The first consist in testing each of the k constraints separately. The intersection of k sub-hypothesis is tested in each point where the distributions are compared. This kind of test, called "intersection-union", is used for example in Bishop et al. (1992). However, Dardanoni et Forcina (1999) and Davidson and Duclos (2000) have demonstrated that this test procedure has relatively low power, since it ignores the covariance structure of the vector $\hat{\boldsymbol{\delta}}$. The second approach, that will be followed here, amounts to simultaneously test the k constraints, using a Wald test and explicitly taking into account the covariance structure of the estimated poverty indices differences vector. The general principle of the test amounts to compare the distances between the estimated vector $\hat{\boldsymbol{\delta}}$ to the sets defining respectively the null

⁴¹In order to use non-stochastic poverty lines, these percentiles are computed from administrative tax records (the Revenus Fiscaux data) that are independent of the BDF survey.

⁴²In our application this formula has been adapted to account for sample weights.

⁴³See Davidson and Duclos (2000), theorem 1. p.1441 for the expression of this matrix.

hypothesis and the alternative hypothesis. Of course, Wald tests can also be used to test the equality of two distributions.

Equality tests The Wald test for the equality of two distributions is relatively easy to implement and resort to a χ^2 test. The null hypothesis is given by $\mathbf{H}_0 : \boldsymbol{\delta} = 0$. One can show (Beach & Davidson 1983, Davidson & Duclos 2000) that under the null hypothesis, the vector $\hat{\boldsymbol{\delta}}$ is asymptotically normal and we have :

$$\hat{\boldsymbol{\delta}} \sim \mathcal{N}\left(0, \frac{\boldsymbol{\Sigma}_A}{N_A} + \frac{\boldsymbol{\Sigma}_B}{N_B}\right)$$

Hence the test statistic T_1 has the following asymptotic distribution under the null hypothesis:

$$T_1 = \hat{\boldsymbol{\delta}}' \left(\frac{\boldsymbol{\Sigma}_A}{N_A} + \frac{\boldsymbol{\Sigma}_B}{N_B} \right)^{-1} \hat{\boldsymbol{\delta}} \sim \chi_k^2$$

Stochastic dominance tests Stochastic dominance tests are more complex to implement since in this case, the set corresponding to the null hypothesis is defined by an inequality constraint. The hypothesis are given by: $\mathbf{H}_0 : \boldsymbol{\delta} \in \mathbb{R}_+^k$ against $\mathbf{H}_1 : \boldsymbol{\delta} \notin \mathbb{R}_+^k$. The Wald test statistic with such constraints has been developed by Kodde and Palm (1986) and Wolak (1989). For this set of hypotheses, the test statistic T_2 is defined by:

$$T_2 = \min_{\boldsymbol{\delta} \in \mathbb{R}_+^k} \|\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}\|$$

with $\|x\| = x' \boldsymbol{\Sigma}^{-1} x$. Kodde and Palm (1986) have shown that the statistic T_2 is distributed as a mixture of χ^2 distributions:

$$T_2 \sim \bar{\chi}^2 = \sum_{j=0}^k w(k, k-j, \boldsymbol{\Sigma}) Pr(\chi_j^2 \geq c)$$

with $w(k, k-j, \boldsymbol{\Sigma})$ the probability that $k-j$ elements of $\boldsymbol{\delta}$ are strictly positive. The distribution of the $\bar{\chi}^2$ distribution have not been tabulated, but lower and upper bounds of critical values are available. When these bounds do not allow to reach a conclusion we estimate the critical values of the statistic T_2 using Monte-Carlo simulation⁴⁴.

⁴⁴We draw 10.000 multivariate normal vectors with mean 0 and covariance matrix $\boldsymbol{\Sigma}$, and compute the proportion of vectors with j positive elements (for $j \in (0, k)$). This proportion is an estimate of the weight $w(k, j, \boldsymbol{\Sigma})$.

B- Summary statistics and results for intermediate waves

Table A-1: Sample summary statistics

	1979	1984	1989	1994	2000
Occupation of the father (%)					
Farmers	21.65	18.23	15.69	12.95	10.55
Artisans	13.93	11.29	12.13	13.61	12.26
H-grade prof.	6.95	6.97	7.92	15.25	15.05
L-grade prof.	7.37	9.63	11.45	8.13	8.52
Non-man. workers	11.18	14.03	12.2	16.28	18.49
Manual workers	38.92	39.85	40.6	33.79	35.13
Obs Imputed	377	233	158	149	0
Mean Income					
before imputation	16182	16428	17071	18276	18180
after imputation	16070	16590	17161	18178	18180
Obs	4231	4428	3529	4644	3984

Notes : mean income in Euros 2002. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers.

Table A-2: Stochastic dominance tests 1984-1994

1984	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	< ₁	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁
1989	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	?	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	> ₁	> ₁
Non-man. workers	-	-	-	-	-	> ₁
1994	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	< ₁	< ₁	< ₁	< ₁	> ₁
Artisans	-	-	< ₁	?	> ₁	> ₁
H-grade prof.	-	-	-	> ₁	> ₁	> ₁
L-grade prof.	-	-	-	-	=	> ₁
Non-man. workers	-	-	-	-	-	> ₁

Notes : The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers. Each element in the table indicates the result of the comparison of the income distribution of the groups in row and column using the tests presented in section 4.2 :

>_{*i*}: the row dominates the column for order *i* stochastic dominance ;

<_{*i*}: the column dominates the row for order *i* stochastic dominance;

=: the distributions are equal;

?: the distributions cannot be ranked using first and second order stochastic dominance.

Table A-3: Lorenz dominance tests 1984-1994

1984	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	=	=	<	<	<
Artisans	-	-	=	?	=	=
H-grade prof.	-	-	-	=	=	=
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=
1989	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	=	=	=	?	<
Artisans	-	-	=	=	?	?
H-grade prof.	-	-	-	=	?	?
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	?
1994	Farmers	Artisans	H-grade prof.	L-grade prof.	Non-man. workers	Manual workers
Farmers	-	=	=	<	<	<
Artisans	-	-	=	<	<	<
H-grade prof.	-	-	-	?	<	<
L-grade prof.	-	-	-	-	=	=
Non-man. workers	-	-	-	-	-	=

Notes : The occupational group refers to social origin. H-grade prof. : higher-grade professionals. L-grade prof. : lower-grade professionals. Non-man. workers : non-manual workers. Each element in the table indicates the result of the comparison of the Lorenz curves of the groups in row and column using the criterion of Lorenz dominance :

$>_i$: the row dominates the column for Lorenz dominance ;

$<_i$: the column dominates the row for Lorenz dominance;

=: the Lorenz curves are equal;

?: the Lorenz curves cannot be ranked using Lorenz dominance.