Working Papers / Documents de travail

The Effect of Aspirations on Inequality: Evidence from the German Reunification using Bayesian Growth Incidence Curves

Edwin Fourrier-Nicolai
Michel Lubrano

WP 2019 - Nr 14

The Effect of Aspirations on Inequality: Evidence from the German Reunification using Bayesian Growth Incidence Curves

Edwin Fourrier-Nicolai[*] Michel Lubrano[†]

February 2019

Abstract

A long-standing literature has investigated the formation of aspirations and how they shape human behaviours but a recent interest has been devoted on the interplay between aspirations and inequality. Because aspirations are socially determined, household investment decisions tend to be reproduced according to the social context which fosters inequality to persist. We empirically examine the role of aspirations on inequality using a natural experiment. We exploit an exogenous variation of social aspirations determined by the exposure to Western German TV broadcasts in the GDR before the reunification. We measure the treatment effect on wage inequality by comparing inequality changes between the treatment and the control regions after reunification. We use an heteroskedastic parametric model for income with a treatment effect and sample selection into the labour market. We derive analytical formulae for the growth incidence curve of Ravallion and Chen (2003) and poverty growth curve of Son (2004) for the log-normal distribution. Based on those curves, we provide Bayesian inference and a set of tests related to stochastic dominance criteria. We find evidences that aspirations - through exposure to Western German broadcasts - have significantly affected inequality. We find that this effect was detrimental in terms of inequality and poverty. However, we cannot conclude about the persistence of the effect after 1995.

Keywords: Inequality, Social aspirations, Bayesian inference, Treatment effect
JEL codes: D31, D91, C11, C21

[*]Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE; Email: edwin.fourrier@univ-amu.fr; Address: 5 boulevard Bourdet, 13001 Marseille, France.

[†]School of Economics, Jiangxi University of Finance and Economics and Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE, Marseille, France. Email: michel.lubrano@univ-amu.fr. Amse-Greqam, 5-9 Bd Maurice Bourdet, 13001 Marseille, France.

1 Introduction

The fall of the Berlin Wall is one of these ambiguous bittersweet events that has marked definitively the end of the Cold War. On the one hand, German reunification and recovered freedom were celebrated. But on the other hand, the whole East-German society was devastated. This ambiguous sentiment was crystallized for instance in the popular neologism *ostalgie* which refers to the nostalgia for the communist era in former Eastern European countries like East-Germany. Numerous novels and films (e.g. "Good Bye, Lenin!") have illustrated the brutal changes that East-Germans experienced at that time. Nostalgia can be the result of disillusioned expectations and aspirations during the years following the German reunification. In this paper, we study how differentials in aspirations have shaped inequality after the German reunification.

A recent literature has investigated the formation of aspirations and how they shape human behaviours. Particularly, the literature abounds of empirical evidences suggesting that aspirations are influenced both by individual circumstances and by social interactions. Among others,[1] Stutzer (2004) and Knight and Gunatilaka (2012) are illuminating in this respect as they find that income aspirations increase with people's income as well as with the average income in the community they live in. Genicot and Ray (2017) proposes a general definition of aspirations as income or wealth thresholds that enter into the individual utility functions as reference points and which depends on both individual achievements and socio-economic contextual outcomes. Because aspirations are socially determined, household investment decisions tend to be reproduced according to the social context which fosters poverty and inequality to persist. Because of the evident difficulty to observe aspirations, there is no empirical evidence that aspirations do affect inequality.

We empirically examine the causal role of aspirations on income inequality using a natural experiment. In this attempt, we exploit an exogenous variation of social aspirations determined by exposure to West-German television broadcasts in East-Germany before reunification. While most East-Germans could (and, according to all available evidence, enthusiastically did) watch West-German television channels, West-German broadcasts did not reach the inhabitants of some regions of East-Germany, sometimes called the "valley of the innocent" (Hyll and Schneider 2013). This natural experiment has been recently used, for instance, in Hyll and Schneider 2013 and Bursztyn and Cantoni (2016) to show that exposure to West-German television broadcasts has affected material aspirations and consumption. Therefore, we ask whether aspirations - through exposure to West-German television broadcasts - have affected not only the mean income but also the variance of the income distribution. We measure the treatment effect on income inequality by comparing distributional changes between the treatment and the control groups after reunification using Bayesian inference. Particularly, we measure distributional changes using the growth incidence curve of Ravallion and Chen (2003) and poverty growth curve of Son (2004) in a parametric framework. Those curves allows us to quantify (in terms of mean and inequality) how the income distribution of the treatment group has diverged from that of the control group because of the exposure to West-German television broadcasts. The advantage of this ap-

[1] For instance, Luttmer (2005) and Clark and Senik (2010) provide empirical evidences of an effect of income comparison on own well-being.

proach is that both curves are closely related to stochastic dominance as shown in Duclos (2009) and Araar et al. (2009). Consequently, distributional changes measured using these curves are consistent with a large set of utility functions as well as a large set of poverty and inequality indices. A growing literature make use of these curves to assess distributional changes, see Lakner and Milanovic (2016) using the world income distribution, Chancel and Piketty (2017) for the Indian income distribution and Novokmet et al. (2017) for Russia among many other examples.

We propose a parametric model of income formation allowing for observable differences between the treatment and control groups. A parametric approach is particularly suitable for various reasons. Indeed, a parameter-free approach is quite demanding in terms of sample size whereas the amount of data available for the researcher is limited in most of empirical studies. Moreover, studying inequality often requires to deal with trimming data problems which are soften by the parametric assumption. Furthermore, the parametric approach allows to have a tractable expression of the distributional changes occurring between two periods and to derive statistical inference for inequality changes directly from the parameters using Bayesian methods. We assume that the income distribution follows a log-normal process; or equivalently, that the log-income follows a normal process. The log-normal process is a very convenient way to model middle-sized incomes as discussed for instance in Aitchison and Brown (1957) (see also Anderson et al. 2014 in the context of a mixture model with two groups). The advantage of adopting a parametric framework is that we can easily control for observed and unobserved heterogeneity using standard methods, but also in that we can derive analytical forms for the growth incidence curve of Ravallion and Chen (2003) and the poverty growth curve of Son (2004). More precisely, the growth incidence curve is a function of the parameters of the wage equation. We use an heteroskedastic model of the log-income where the treatment can affect both the mean and the variance of the log income. Because distributional changes are measurable only on positive incomes and because individuals with positive incomes may not be a random sample of the population, we control for potential sample selection following the Heckman (1979)'s approach. Based on those curves, we provide Bayesian inference and a set of tests related to stochastic dominance criteria. Particularly, we test if a distributional change has been welfare improving in terms of first-order (second-order) stochastic dominance, if a distributional change has been relatively pro-poor, if every quantile has benefited equally from growth and finally, if the distributional change of a group is preferred to that of the other group.

Our approach has certain limits. Both curves measure changes in the (anonymous) income distribution, ignoring income mobility.[2] In other terms, we look at changes in the mean income of each quantile, but we ignore who are in a particular quantile. An increase in the mean income of the median group does not imply that all the people at the median in the first period are wealthier in the second period since there might have been social mobility in between. Therefore, the growth incidence curve and the pro-poor growth curve can be estimated using repeated cross-

[2]Jenkins and Van Kerm (2006) were the first to consider re-ranking using a decomposition of the variation of a generalized Gini index between progressivity and re-ranking with concentration curves. They analyse changes in income inequality in the US and contrast these changes with the evolution of Germany. Grimm (2007) has shown that different conclusions can be reached if removing the axiom of anonymity when computing growth incidence curves. Bourguignon (2011) builds on this approach and provides dominance criterion.

section data instead of balanced panel data. Using the SOEP, we find evidences that aspirations - through exposure to Western German broadcasts - have significantly affected inequality in post-reunification East-Germany. We find that this effect was detrimental in terms of inequality and poverty. However, we cannot conclude about the persistence of the effect after 1995. This paper contributes to the literature on the social determinants of poverty and inequality. In doing so, we provide a method to measure the effect of a treatment on a whole distribution. We contribute to the literature on the econometrics of poverty and inequality by providing parametric formulae for the growth incidence curve and the pro-poor growth curve for the log-normal distribution. We also provide Bayesian inference to test for numerous distributional changes related to welfare criteria.

The paper is organized as follows. In section 2, we present the conceptual framework, the natural experiment, the data and the empirical methodology. In section 3, we explain how growth incidence curves and related measures are derived. We also explain the relationship between these curves and stochastic dominance. In section 4, we provide parametric growth incidence curve and pro-poor growth curve for the log-normal distribution with functional heteroskedasticity and potential sample selection. In section 5, we provide Bayesian inference for the curves and propose a set of tests based on stochastic dominance criteria. In section 6, we provide the empirical results and we carry out a counter-factual exercise to gauge the effect of aspirations on inequality. The last section concludes.

2 Aspirations in East-Germany before reunification

2.1 Conceptual framework

A growing literature has been investigating the formation of aspirations and how they can affect socio-economic outcomes. Particularly, the literature abounds of empirical evidences suggesting that aspirations are influenced both by individual circumstances and by social interactions. Among others,[3] Stutzer (2004) and Knight and Gunatilaka (2012) are illuminating in this respect as they find that income aspirations increase with people's income as well as with the average income in the community they live in. Genicot and Ray (2017) proposes a general definition of aspirations as income or wealth thresholds that enter into the individual utility functions as reference points and which depends on both individual achievements and socio-economic contextual outcomes.

On the other hand, a recent empirical literature has been investigating how television can shape aspirations. In the context of the German reunification, Hyll and Schneider (2013) find consistent evidences that watching Western television in the former GDR increased material aspirations substantially while Bursztyn and Cantoni (2016) find that Western television exposure affects the structure of consumption. In doing so, television affects the perceived socio-economic context or, in other terms, the reference group of individuals; that is, income and consumption comparisons take place not only with individual's actual reference group (e.g. relatives, friends,

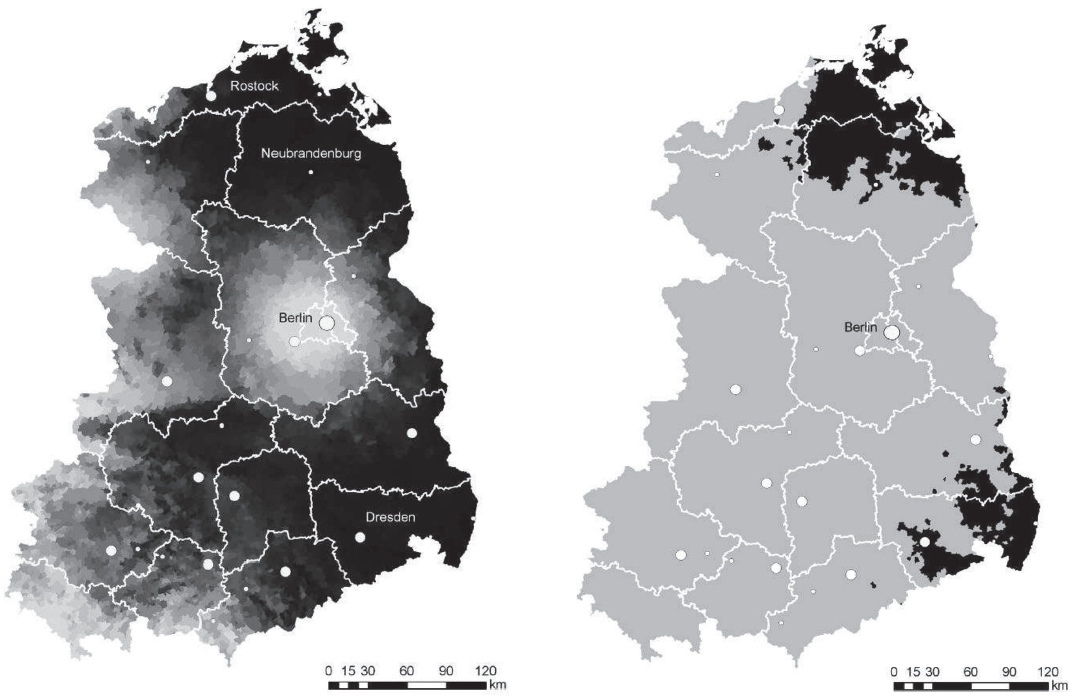
[3]For instance, Luttmer (2005) and Clark and Senik (2010) provide empirical evidences of an effect of income comparison on own well-being.

neighbours, and colleagues), but also with a virtual reference group consisting of television characters. In this attempt, advertising and entertainment programs play a non-negligible role in shaping mental representations (e.g. Chong and Ferrara 2009 or Bursztyn and Cantoni 2016). Because consumption signals one's economic status to others and since individuals have a demand for prestige then this virtual reference group affects consumption behaviours. This phenomenon gave rise to a prolific literature on conspicuous consumption initiated with the seminal work of Veblen (1899). Of course, television exposure may also affect material conditions indirectly through its effect on other living conditions e.g. women status (Jensen and Oster 2009), divorce (Chong and Ferrara 2009), fertility (Boenisch and Hyll 2015) and criminality (Friehe et al. 2017) for example.

Even if the effect of aspirations has been studied for a wide range of socio-economic outcomes, their effect on inequality has been largely ignored in the empirical literature while Genicot and Ray (2017) demonstrates theoretically that there is an interplay between aspirations and wealth inequality. More precisely, they show in a model of socially determined aspirations that economy-wide outcomes determine individual aspirations, which in turn determine investment decisions and then social outcomes. Particularly, when aspirations are socially monotone (non-decreasing in society-wide incomes), the initial income distribution will diverge (i.e. increase of between-inequality and decrease of within-inequality). The only case where the initial distribution would converge towards the equal distribution is when the initial distribution has already a large degree of equality. As a consequence, aspirations may create social poverty traps increasing the persistence of wealth inequality and poverty. Therefore, we focus mainly on the effect of aspirations on income inequality and more particularly on labour earnings as they are likely to be influenced by aspirations more than other sources of income.

2.2 The natural experiment

We discussed to what extent watching television might affect viewers' aspirations. However, it is also conceivable that individuals with high aspirations are more likely to watch television, especially if television programs depict wealthier comparison groups. Conversely, it is possible that individuals with low aspirations watch television excessively simply for entertainment owing to low opportunity costs (Frey et al. 2007). It is therefore difficult to measure the causal effect of television exposure on aspirations which, in turn, would affect inequality. We address the problem of endogeneity by exploiting a natural experiment occurring in the Eastern part of the divided Germany. Owing to exogenous variations in the signal strength we can identify two comparable groups with different access to television broadcasting. While one group (the control group) could only receive East German television channels, the other group (the treatment group) was exposed to West German television channels too. More precisely, even if most East-German households could receive West-German television before the reunification, there remained locations that could not receive West-German television broadcasts due to geographical obstacles. Bursztyn and Cantoni (2016) estimate the signal strength of Western television relying on a signal propagation model that takes the Earth's curvature and elevation features into account. While in the absence of any obstacle, an electromagnetic signal declines in strength with the square of the distance from its source; in practice, geography and topography are the main determinants of actual availability of TV signals. The GDR regions without access were located either in the North-East and in the South-East of the country, and were either too far away from the transmitters or were located on the other side of mountains that blocked the signals. The left panel of Figure 1 provides the estimated signal strength in 1989 by Bursztyn and Cantoni (2016). Because the signal strength is continuous, we have to find the value above which individuals can actually watch Western television programs and under which they actually cannot. Gathering survey data and anecdotal evidences, Bursztyn and Cantoni (2016) set the signal strength in Dresden as the threshold for partitioning the population into treatment and control groups. Then, they consider all municipalities with signal strength equal to or below the Dresden's average signal strength to be in the control area, that is the population non-exposed to Western TV broadcasts. The treatment area comprises all regions with a positive probability of reception of Western television broadcasts. The right panel of Figure 1 represents the control area in black and the treatment area in grey as provided by Bursztyn and Cantoni (2016). We use their results to define a treatment group (those who were exposed to Western TV signals) and a control group (those deprived from Western television signals). About 85% of the GDR population was exposed to Western television broadcasts, while the remaining 15% had access only to East-German television broadcasts.



Signal strength Treatment and control regions

Figure 1: Exposure to Western TV broadcasts in East-Germany (1989)

Black areas represent weak signal areas where the signal strength is weakly lower than in Dresden, dots represent major cities and white lines represent district borders. *Source: Bursztyn and Cantoni (2016), Figures 1 and 3.*

Although the process governing the television exposure resembles random assignment, it may actually not be. As a consequence, we should ask whether the treatment and the control groups were similar and, if not to what extent. All existing empirical evidences suggest that both regions were highly homogeneous with respect to cultural and socio-economic development as shown in Table 1 of Bursztyn and Cantoni (2016) for instance. They compare both regions in 1955 and 1990 and they do not observe any differential trend for the two groups between 1955 (before television became popular in the GDR) and 1990 (during the reunification). In the same direction, Hyll and Schneider (2013) demonstrate that economic, political, and social conditions in the control region did not systematically differ from conditions in the other regions of East-Germany. We also investigate this issue by testing for differences in the individual characteristics of both groups just after the reunification and we confidently reject this hypothesis. Last, there is no evidence that individuals who were more interested in Western broadcasts had moved into an area with better reception before the reunification. Particularly, Bursztyn and Cantoni (2016) argue that the centralization of the economy and the chronic shortage inhibits selective spatial sorting across groups before the reunification.

Similarly, a confounding factor could exist if there were differences in TV set equipments between the treatment and control regions. Most of empirical evidences provided in Bursztyn and Cantoni (2016) go in the direction of no noticeable difference between both regions. Particularly, they argue that the limited availability of Western TV broadcasts did not prompt households in the control region to buy fewer television sets. Television was a basic good and the great majority of households had a television set in East-Germany (98% of households in 1989). In addition, they report that 46% had a colour television set and one out of six households had more than one television set, leading to an average of 117 television sets per 100 households. The two production facilities for television sets in East-Germany were located in Dresden (in the control region) and in Stassfurt (in the treatment region).

Remains the questions of whether households in the exposed region actually consume Western television broadcasts. Most of empirical evidences suggest that the great majority of those who were exposed to Western television actually consumed it. Although watching Western television was officially forbidden by the GDR state, a survey led in 1985 among young people by the *Zentralinstitut fur Jugendforschung* reports that respondents in the treatment region watched more than two hours of West-German television per day on average. Furthermore, Figure 2 of Bursztyn and Cantoni (2016) indicates that 66% of respondents in districts with access to Western television declared they watched Western TV stations daily. In contrast, only 5% of the respondents in the district of Dresden declared so.

2.3 Data

We use the German Socio-Economic Panel (SOEP), a longitudinal representative survey provided by the German Institute for Economic Research (DIW). The survey started in 1984 in the FRG and the first wave surveying East-Germans started in 1990 for a restricted set of variables. We use the regional policy regions (ROR, Raumordnungsregionen) wherein the households lived in 1990 to define whether a household belongs to the treatment or to the control group. The control group was formed by the households located in one of the 5 following ROR in 1990: Stralsund-Greifswald, Rostock, Neubrandenburg, Oberlausitz and Dresden. It is worth mentioning that this geographical unit is not enough accurate so as to obtain a perfect replication of the black area in Figure 1. Therefore, the control group is defined in such a way that it includes the complete black area and incidentally some exposed households. This potential misallocation should be nuanced as the signal strength is a continuous variable and that false-control households are likely to actually have weak television signals. In addition, this misallocation would generate an attenuation bias which only strengthens our case if we actually find an effect. We assign households living in the rest of East-Germany to the treatment group. However, we exclude Berlin from the analysis since it is the capital which was part of both East and West parts of Germany. We restrict the sample to the working-age population (between 16 and 65) in 1990. We then build an unbalanced panel formed of all individuals living in East-Germany (excluding Berlin) which were present in the sample in 1990. This is thus a cohort where the same individuals are followed over time. Sample sizes are provided in Table 1. They naturally decrease over time because of attrition.

The different sources of income and numerous socio-economic characteristics have been reported in the SOEP. Unfortunately, labour income is observed in East Germany only after 1992, where changes due to the reunification had already been in operation. However, the income reported in 1992 is the income of the previous year, namely 1991. Still, the initial period should be 1990, taking 1991 as the initial period would potentially attenuate the treatment effect. Therefore, this only strengthens our case if we want to test the existence of a treatment effect. Labour income corresponds to wage earnings before taxes and redistribution. We use a consumer price index to make income comparable over time and between regions. As we focus on labour income, there is a significant proportion of zero income because of unemployment or because of non-participation onto the labour market. That proportion is lower in East Germany wherein female labour participation was higher. But this spread between East and West Germany is reducing over time. In order to investigate this issue, we transform our data into a balanced panel, then we can shed light on some of the differences that exist between the two groups. Particularly, Table 2 provides transition probabilities of having a zero labour income for each group between 1992 and 1995. Clearly, Table 2 suggest substantive between-group differences in the probability of being excluded from the labour market. We investigate this issue by taking into account the potential sample selection and we allow the treatment to affect the decision to participate onto the labour market. Similarly, we observe significant between-group difference in the probability of migrating to West Germany after the reunification.

For those participating to the labour market, we draw a portray of the labour income distributions for each group after the reunification in Table 3. Particularly, we use the median as a measure of centrality, the Gini coefficient as a measure of inequality and the poverty rate is computed as the proportion of individuals with an income below 60% of the median West-German labour income. Clearly, the poverty rate has decreased over the period with a greater speed in the treatment group whereas inequality has increased more in the treatment group than in the control group. We are particularly interested in determining whether these differences between the treatment and the control groups are due to differences in group characteristics or instead to the effect of television exposure.

	1990	1991	1992	1993	1994	1995
			Sample size			
Total	13 046	12 247	11 702	11 183	10 608	10 100
West-Germany	8 889	8 429	8 085	7 766	7 339	6 975
East-Germany	4 157	3 818	3 617	3 417	3 269	3 125
Treatment group	2 882	2 667	2 506	2 383	2 281	2 188
Control group	771	719	669	622	609	571
			Percentage of zero labour incomes			
Total	-	-	0.220	0.237	0.253	0.270
West-Germany	0.270	0.255	0.252	0.259	0.275	0.288
East-Germany	-	-	0.149	0.188	0.205	0.230
Treatment group	-	-	0.154	0.191	0.213	0.234
Control group	-	-	0.136	0.199	0.195	0.238

Table 1: Sample sizes

for zero labour incomes between 1992 and 1995

	Treatment		Control	
	Zero	Positive	Zero	Positive
Zero	0.715	0.285	0.667	0.333
Positive	0.150	0.850	0.174	0.826

Table 2: Transition probabilities

	1992	1993	1994	1995
	East-Germany			
Median	16 108	16 778	17 849	18 739
Gini	0.308	0.323	0.327	0.319
Poverty rate	0.463	0.440	0.422	0.383
	Treatment group			
Median	16 049	16 778	17 849	18 679
Gini	0.306	0.321	0.322	0.319
Poverty rate	0.464	0.442	0.415	0.380
	Control group			
Median	16 423	16 767	17 836	19 098
Gini	0.315	0.329	0.345	0.322
Poverty rate	0.458	0.436	0.449	0.393
West poverty line	15 280	15 478	15 740	15 683

Zero incomes were excluded from the sample. The poverty line was chosen as 60% of median West-German labour income. Berlin is excluded from all groups.

Table 3: Labour income characteristics for treated and control groups

2.4 Empirical strategy

We ask whether aspirations through exposure to Western television broadcasts has affected income inequality in Germany after the reunification. In doing so, we basically want to assess which of the income distributional changes are attributable to the exposure to Western television broadcasts. This raises a set of empirical challenges.

The growth incidence curve of Ravallion and Chen (2003) is a convenient way of comparing changes in the income distribution between two points in time. It measures how growth is distributed over the quantiles. One advantage of this approach is that it is related to stochastic dominance as shown for instance in Duclos (2009) and Araar et al. (2009). However, it is not clear whether between group differences in distributional changes arise because of the treatment or because of changes in the features of the groups. Consequently, it is not sufficient to compare the growth incidence curve of the treatment with that of the control group. Instead, we should take into account potential changes in the composition of each group between the two periods. Therefore, we want to identify which quantiles of the income distribution have been significantly affected by the treatment taking into account for potential heterogeneity.

A parametric approach is particularly suitable to answer this question for various reasons. Indeed, a parameter-free approach is quite demanding in terms of sample size whereas the amount of data available for the researcher is limited in most of empirical studies. In our case, we have 669 observations in 1995 and only 578 non-zero incomes. Moreover, studying inequality often requires to deal with trimming data problems which are soften by the parametric assumption. Furthermore, the parametric approach allows to have a tractable expression of the distributional changes occurring between two periods and to derive statistical inference for inequality changes directly from the parameters using Bayesian methods. We assume that the income distribution follows a log-normal process; or equivalently, that the log-income follows a normal process. The log-normal process is a very convenient way to model middle-sized incomes as discussed for instance in Aitchison and Brown (1957) or in Anderson et al. 2014 in the context of a mixture of log-normal densities. The advantage of adopting a parametric framework is that we can easily control for observed and unobserved heterogeneity using standard methods, but also that we can derive analytical forms for the growth incidence curve of Ravallion and Chen (2003). More precisely, the growth incidence curve can be expressed as a function of the parameters of the wage equation. We introduce the treatment effect and observed heterogeneity by means of covariates for explaining both the mean and the variance of the log-income while recognizing that these two groups belong to the same population. Because, individuals with positive incomes may not be a random sample of the population, we control for potential sample selection of positive incomes following the Heckman (1979)'s approach.

Numerous hypotheses on distributional changes can be performed using Bayesian inference. Particularly, we test if a distributional change has been welfare improving in terms of stochastic dominance, if a distributional change has been relatively pro-poor, if every quantile has benefited equally from growth and finally, if the distributional change of a group is preferred to that of the other group. Because of the relationship between the growth incidence curve and stochastic dominance at the first-order, our results are consistent with a large set of utility functions and most of poverty and inequality indices. When, one cannot conclude about stochastic dominance at the first-order, we consider stochastic dominance at the second-order using the poverty growth curve of Son (2004).

3 Measurement of inequality changes and stochastic dominance

3.1 Quantile functions and Lorenz curves

Let Y be a continuous random variable (e.g. income) with cumulative distribution function (cdf) $F(y)$ and probability density function (pdf) $f(y)$ with support contained on the non-negative real line. The quantile function is defined as the inverse of the cdf:[4]

$$Q(p) = F^{-1}(y), \qquad \text{or} \qquad Q(p) = \inf_{y \geq 0}(F(y) \geq p).$$

For computing the mean, a poverty or an inequality index, we can use directly the pdf of the random variable Y. For instance the mean is defined as:

$$\bar{y} = \int_0^{\infty} y f(y)\, dy. \tag{1}$$

However, we can use the dual estimator based on the quantile function. Let us consider the change of variable $y = F^{-1}(p)$ and apply it to (1), we get:

$$\bar{y} = \int_0^1 F^{-1}(p)\, dp = \int_0^1 Q(p)\, dp. \tag{2}$$

Let us now consider the Lorenz curve, a widely used measure of inequality introduced in Lorenz (1905). It was originally defined by:

$$\begin{aligned} L(p) &= \frac{1}{\bar{y}} \int_0^y t f(t)\, dt \\ p &= F(y). \end{aligned}$$

Using the same change of variable $y = Q(p)$, Gastwirth (1971) provides the following form of the Lorenz curve:

$$L(p) = \frac{1}{\bar{y}} \int_0^p Q(t)\, dt,$$

which immediately relates the quantile function to the Lorenz curve with:[5]

$$Q(p) = \bar{y} L'(p). \tag{3}$$

[4]The empirical income distribution is formed by n observations of Y, noted y and arranged by increasing order. The sequence of order statistics is noted $y_{[i]}$. The graph of the empirical quantile function is obtained by plotting the n component vector $[p_i = i/n]$ in $[0, 1]$ against the n order statistics. If we normalize this graph by the mean, we get the well-known Pen's parade.

[5]From this expression it becomes clear that the mean income in the population is found at the percentile at which the slope of $L(p)$ (i.e. $L'(p)$) is equal to 1.

A variant of the Lorenz curve representing inequality while taking into account the level of income has been introduced formally in Shorrocks (1983), that is the generalized Lorenz curve. It is simply obtained by multiplying the Lorenz curve by the mean income $\bar{y}$:

$$GL(p) = \int_0^p Q(t)\, dt = \bar{y}L(p).$$

Numerous inequality and poverty measures[6] also rely on the quantile function, and thus, can be derived from the Lorenz curve too (see e.g. Foster and Shorrocks 1988).

3.2 Inequality dynamics

Let us now consider two dates t and $t-1$ and their respective distributions $F_t(y)$ and $F_{t-1}(y)$. The growth incidence curve (GIC), introduced in Ravallion and Chen (2003), measures the growth rate of the p-quantile for every p:

$$g_t(p) = \frac{Q_t(p)}{Q_{t-1}(p)} - 1 \simeq \log Q_t(p) - \log Q_{t-1}(p). \tag{4}$$

Using (3), the GIC can be immediately related to the Lorenz curve with:

$$g_t(p) = \frac{L_t'(p)}{L_{t-1}'(p)}(\gamma_t + 1) - 1 \simeq \log GL_t'(p) - \log GL_{t-1}'(p), \tag{5}$$

where $\gamma_t = (\bar{y}_t - \bar{y}_{t-1})/\bar{y}_{t-1} \simeq \log(\bar{y}_t) - \log(\bar{y}_{t-1})$ is the average growth rate. Two immediate properties can be derived from (5): if inequality does not change then $g_t(p) = \gamma_t$ for all p, and the p-quantile increases if $g_t(p) > 0$.

Thus the growth incidence curve corresponds to the variation of the first derivative of the generalized Lorenz curve. Graphically, the GIC associates the growth rate of income with respect to proportion p of individuals ordered by increasing income. By drawing the horizontal line corresponding to the rate of growth of the mean income (or the median income), the quantiles below that line have a rate of growth of their income which is lower than the growth rate of the mean income (or the median income). Remark that computing the GIC requires only cross-section data at two different times and not longitudinal data. By doing so, we consider only the shapes of the distribution and not individuals destinies per se. This is why the GIC curve is sometimes referred as being anonymous.

An alternative approach for assessing distributional changes has been proposed by Son (2004) who introduces the poverty growth curve (PGC). The initial question of Son (2004) was to determine whether the mean income of the lower quantiles (corresponding to the poor) is growing quicker than the mean income of the other quantiles. The poverty growth curve is defined as the variation in percentage of the average income of the bottom $p\%$ of the population and corresponds to $\Delta \log(\bar{y}_p)$, where $\bar{y}_p$ is the average income of the bottom $p\%$. Using (2), the Lorenz curve can be written as:

$$L(p) = \frac{\int_0^p Q(t)\, dt}{\int_0^1 Q(t)\, dt} = \frac{p\bar{y}_p}{\bar{y}},$$

[6] Among others, the Gini coefficient, the FGT indices of Foster et al. (1984) and the TIP curve as shown in Fourrier-Nicolai and Lubrano (2017) for instance.

the poverty growth curve corresponds to:

$$G_t(p) = \Delta \log \bar{y}_p = \Delta \log \bar{y} + \Delta \log L(p) = \Delta \log GL(p). \tag{6}$$

In this context, growth is pro-poor if the variations of the Lorenz curve are positive for all p up to a given value. The PGC is thus equal to the variation of the generalized Lorenz curve. Because $L(1) = 1$ then $\Delta L(1) = 0$ and the PGC is equal to γ_t at $p = 1$. Growth is pro-poor when $G_t(p) > \gamma_t$, which means that the $G_t(p)$ curve is decreasing in p as $G_t(1) = \gamma_t$. Poverty simply decreases when $G_t(p) > 0$ for all $p < 1$. When $0 < G_t(p) < \gamma_t$ for all $p < 1$, there is a phenomenon of trickle-down growth, that is to say poverty is reduced but not as much as it could because the rich are receiving proportionally more.

Using the same notation and approximation, the GIC can be written as:

$$g_t(p) = \Delta \log Q(p) = \Delta \log \bar{y} + \Delta \log L'(p) = \Delta \log GL'(p),$$

so that the two curves can be compared. Both measures are obtained as the variation of the log of the mean income plus the variation of either the log of the Lorenz curve or the log of its derivative. While the growth rate of income at the p-quantile is used for the GIC, the PGC is based on the estimation of the growth rate of the mean income up to the p-quantile.

3.3 Stochastic dominance

Remains the question of whether the distributional changes have been favourable for the economy. Under the veil of ignorance, stochastic dominance is used to determine whether a distribution is preferred to another. Particularly, if a distribution first-order dominates another one, then the first distribution is preferred for any non-decreasing social utility function (Atkinson 1970). Let us consider an income distribution growing between two periods with y_{t-1} and y_t with a growth rate γ and a common poverty line z. First-order stochastic dominance of y_{t-1} by y_t up to a poverty line z implies that $F(y_{t-1}) \geq F(y_t)$ for all $y \leq z$. This means that the proportion of individuals below the poverty line is always greater in $F(y_{t-1})$ than in $F(y_t)$, for any poverty line lower than z. Since stochastic dominance is essentially a comparison of the cumulative distribution functions, the quantile functions are related to stochastic dominance as well, this is the p-approach to dominance of Davidson and Duclos (2000). Thus, first-order stochastic dominance of y_{t-1} by y_t implies that $F^{-1}(y_{t-1}) \leq F^{-1}(y_t)$ for all y. It follows directly from equation (4) that first-order stochastic dominance of y_{t-1} by y_t is equivalent to:

$$g_t(p) > 0, \quad \forall p \in [0, 1].$$

In other terms, first-order stochastic dominance of the first period over the second period is verified if and only if the growth incidence curve is positive for every quantile.

If the growth incidence curve is negative for some values of p, then we cannot conclude unambiguously about whether the distributional changes have been welfare-improving or not. In such case, one has to impose more normative conditions by considering stochastic dominance at the second order. Particularly, if a distribution second-order dominates another one, then the first distribution is preferred for any non-decreasing and concave (risk-averse) social utility function. Consequently, second-order dominance is weaker than the first order stochastic dominance and is more likely to be satisfied. Since generalised Lorenz dominance is strictly equivalent to second-order stochastic dominance (Atkinson 1987, Foster and Shorrocks 1988), it follows directly from equation (6) that second-order stochastic dominance of y_{t-1} by y_t is equivalent to:

$$G_t(p) > 0, \quad \forall p \in [0, 1[,$$

Therefore, second-order stochastic dominance of the first period over the second period is verified if and only if the poverty growth curve is positive for every quantile.

Duclos (2009) and Araar et al. (2009) go a step further on and state that growth is relatively pro-poor if:

$$g_t(p) > \gamma_t, \quad \forall p \in [0, F(z)].$$

This condition is verified if the quantiles of the poor increase at a pace greater than the average growth. Taking into account inequality among the poor, we can test whether:

$$G_t(p) > \gamma_t, \quad \forall p \in [0, 1[.$$

4 A parametric treatment effect model for comparing inequality dynamics

4.1 Growth incidence curve for the log-normal model

We first detail the simple log-normal model in order to derive the growth incidence and the poverty growth curves. The individual income for a given period can be modelled in a simple way:

$$\log(y_i) = \mu + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2). \tag{7}$$

In this formulation, σ monitors directly inequality as for instance the Gini coefficient depends only on σ in this case:

$$G = 2\Phi(\sigma/\sqrt{2}) - 1,$$

where Φ is the cumulative distribution function of the standard normal distribution. The mean income is a function of the two parameters of the distribution with:

$$\mathrm{E}(y) = \exp(\mu + \sigma^2/2),$$

while the median income is simply equal to $\exp(\mu)$. As the distribution function corresponds to:

$$F_\Lambda = \Phi\left(\frac{\log y - \mu}{\sigma}\right),$$

the quantile function is given by:

$$Q(p) = \exp(\mu + \sigma\Phi^{-1}(p)),$$

and the Lorenz curve corresponds to:

$$L(p) = \Phi(\Phi^{-1}(p) - \sigma).$$

We can now deduce the GIC and the PGC for two log-normal distributions with parameters $(\mu_{t-1}, \sigma_{t-1})$ and (μ_t, σ_t):

$$g_t(p) = (\mu_t + \sigma_t \Phi^{-1}(p)) - (\mu_{t-1} + \sigma_{t-1}\Phi^{-1}(p)), \tag{8}$$

$$G_t(p) = (\mu_t + \sigma_t^2/2) - (\mu_{t-1} + \sigma_{t-1}^2/2) + \log\left(\frac{\Phi(\Phi^{-1}(p) - \sigma_t)}{\Phi(\Phi^{-1}(p) - \sigma_{t-1})}\right). \tag{9}$$

The mean growth rate γ_t and the median growth rate $\bar{\gamma}_t$ are given by:

$$\gamma_t = \mu_t - \mu_{t-1} + \frac{\sigma_t^2 - \sigma_{t-1}^2}{2}, \qquad \bar{\gamma}_t = \mu_t - \mu_{t-1}.$$

Clearly, the shape of the growth incidence curve is entirely determined by the cdf of the standard normal distribution. As for $p = 0.5$, the standardized quantile function is equal to zero, then $g_t(0.5) = \bar{\gamma}_t$ and when varying σ_{t-1} and σ_t the GIC is turning around this point.

4.2 Modelling the decision to participate

It is common to observe a substantial proportion of zeros in the income distribution which exacerbates when studying income before taxes and transfers. As shown in Table 1, our final sample contains 15% of zero labour incomes in 1992, a number which goes up to 23% in 1995. These zero incomes corresponds to the individuals who have decided not to participate onto the labour market, presumably because they had a reservation wage which was higher than their potential market wage. Individuals with an offered wage below their reservation wage will not work. If the treatment has a positive influence on potential market wages, people in the control group will have, on average, a lower offered wage and therefore a lower employment rate than the treatment group. As a consequence, one will only observe the wages of individuals in the control group who receive comparatively high wage offers. This introduces a common sample selection bias wherein the sample is not representative of the population.

Heckman (1979) has proposed a simple practical solution for such situations, which treats the selection problem as an omitted variable problem. This method has the advantage of being easily implementable and robust to many sample selection processes (Puhani 2000). As in our sample, we observe both positive and zero wages, we can define a dummy variable P_{it} which equals one when the observed wage is positive and zero otherwise. We consider the utility of participating P_{it}^* as being determined by a probit model with underlying equation:

$$P_i^* = \tilde{z}_i'\zeta + u_i \qquad u_i \sim N(0,1), \tag{10}$$

and an observation rule:

$$P_{it} = \begin{cases} 1 & \text{if } P_i^* > 0, \\ 0 & \text{if } P_i^* \leq 0. \end{cases} \tag{11}$$

As shown in Cameron and Trivedi (2005, page 541), it follows that:

$$\mathrm{E}(P_i^*|P_i^* > 0) = \mathrm{E}(u_i > -\tilde{z}_i'\zeta) = \lambda(\tilde{z}_i'\zeta),$$

where $\lambda(\cdot)$ is the well-known inverse Mills ratio and is defined as:

$$\lambda(\tilde{z}_i'\zeta) = \frac{\phi(\tilde{z}_i'\zeta)}{\Phi(\tilde{z}_i'\zeta)}.$$

The corresponding truncated variance is obtained as

$$\mathrm{Var}(P_i^*|P_i^* > 0) = 1 - \lambda(\tilde{z}_i'\zeta)\,\tilde{z}_i'\zeta - \lambda^2(\tilde{z}_i'\zeta).$$

4.3 A wage equation with heteroskedastic errors

In the absence of zero income, the parametric wage equation with observed heterogeneity would be:

$$\log(y_i) = \mu + \tilde{x}_i'\beta + \epsilon_i. \tag{12}$$

However, we have to take into account the fact that we are in fact in a bivariate model (decision to participate and wage equation) with potential correlated error terms:

$$\begin{pmatrix} u_i \\ \epsilon_i \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}\right)$$

So in fact we are interested in modelling the conditional expectation:

$$\mathrm{E}(\log(y_i)|P_{it}^* > 0) = \mu + \tilde{x}_i'\beta + \mathrm{E}(\epsilon_i|u_i > -\tilde{z}_i'\zeta).$$

This expectation is computed in Cameron and Trivedi (2005, page 549), using arguments related to the conditional normal distribution. The regression equation we are interested in is:

$$\log(y_i) = \mu + \tilde{x}_i'\beta + \rho\lambda(\tilde{z}_i'\zeta) + v_i,$$

with heteroskedatic errors of the form:

$$\mathrm{Var}(\log(y_i)|P_i^* > 0) = \sigma_2^2 - \rho(\tilde{z}_i'\zeta\lambda(\tilde{z}_i'\zeta) + \lambda^2(\tilde{z}_i'\zeta)).$$

This model is correctly identified provided there is at least some variables which are not common between $\tilde{z}_i$ and $\tilde{x}_i$.

4.4 GIC and PGC with treatment effect

Equation (12) makes explicit that observed characteristics x_i may affect the median income as compared to equation (7). This is a necessary assumption if we want to have $g_t(p)$ and $G_t(p)$ to depend on the treatment status and the individual characteristics as well. Since, the treatment may have affected the income distribution directly through inequality, we introduce heterogeneity at the level of σ as well. Thus, we have a regression model with functional heteroskedasticity:

$$\begin{aligned}\log(y_i) &= \mu + \tilde{x}_i'\beta_1 + \beta_2\lambda(\tilde{z}_i'\zeta) + v_i, \\ &= \mu + x_i'\beta + v_i, \qquad v_i \sim N(0, \sigma_i^2),\end{aligned} \tag{13}$$

where x_i contains both $\tilde{x}_i$ and $\lambda(\tilde{z}_i'\zeta)$. The functional heteroskedasticity arising from sample selection is given by:

$$\sigma_i^2 = \sigma^2(1 - \rho(\tilde{z}_i'\zeta\lambda(\tilde{z}_i'\zeta) + \lambda^2(\tilde{z}_i'\zeta)) = \sigma^2 h(z_i'\delta). \tag{14}$$

This form of heteroskedasticity is fairly constrained. We would like to take into account a more general form of heteroskedasticty which would occur even in the absence of sample selection or more simply even if $\rho = 0$. So we consider the general form $h(z_i'\delta)$ where z_i represents a set of exogeneous variables including both elements of x_i and $\tilde{z}_i$ to which we add $\lambda(\tilde{z}_i'\zeta)$ and its square.

Let us now compute the conditional mean of y as:

$$\mathrm{E}(y|x_i) = \exp(\mu + x_i'\beta + \sigma^2 h(z_i'\delta)/2), \tag{15}$$

and the conditional quantile function:

$$Q(p) = \exp(\mu + x_i'\beta + \sigma\sqrt{h(z_i'\delta)}\Phi^{-1}(p)). \tag{16}$$

The conditional Lorenz curve is:

$$L(p) = \Phi(\Phi^{-1}(p) - \sigma\sqrt{h(z_i'\delta)}).$$

The GIC and PGC formulae we are interested in are simple translations of the formulae derived for the simple log-normal case:

$$\begin{aligned}g_t(p) &= \mu_t + x_i'\beta_t + \sigma_t\sqrt{h(z_i'\delta_t)}\Phi^{-1}(p) \\ &\quad -(\mu_{t-1} + x_i'\beta_{t-1} + \sigma_{t-1}\sqrt{h(z_i'\delta_{t-1})}\Phi^{-1}(p)),\end{aligned} \tag{17}$$

$$\begin{aligned}G_t(p) &= \mu_t + x_i'\beta_t + \frac{\sigma_t^2}{2}h(z_i'\delta_t) - (\mu_{t-1} + x_i'\beta_{t-1} + \frac{\sigma_{t-1}^2}{2}h(z_i'\delta_{t-1})) \\ &\quad + \log\left(\frac{\Phi(\Phi^{-1}(p) - \sigma_t\sqrt{h(z_i'\delta_t)})}{\Phi(\Phi^{-1}(p) - \sigma_{t-1}\sqrt{h(z_i'\delta_{t-1})})}\right).\end{aligned} \tag{18}$$

Consequently, we have relaxed constraints on the shapes of $g_t(p)$ and $G_t(p)$ as both μ_i and σ_i are varying, we have removed the previous constraint of symmetry and the two curves can also be moved by translation. Remark also that both curves depend now on the value of the observed characteristics x_i. Therefore, computation of the curves requires to select a particular value of the x_i. Particularly, we can compute the global GIC and PGC at the average values of the characteristics $\bar{x}$ and $\bar{z}$ for the whole sample whereas the treatment and control curves can be computed using the average characteristics of the treatment and control groups respectively while assuming the parameter values to be the same. The GIC for the two subgroups will be different as long as the two subgroups have different sample characteristics, so we control for observed heterogeneity between the groups.

5 Bayesian inference

We can first conduct inference of the selection equation. It relies essentially on a Gibbs sampler if we follow Koop (2003, pages 214-216), but a Metropolis-Hastings algorithm is also possible (see e.g. Marin and Robert 2007). Then knowing the posterior mean of ζ, we can compute $\lambda(\tilde{z}_i'\zeta)$. Conditionally on this value, we can conduct inference on the heteroskedastic model in a second step, which implies a Metropolis generator for δ, the parameter of the skedastic function.

5.1 The probit model

Bayesian inference for the probit model was first proposed by Albert and Chib (1993) using a Gibbs sampler. Following equation (10), if the latent variable P_i^* were known, this model would be a linear regression model with unit variance. Under a non-informative prior, the posterior density of ζ conditionally on P_i^* and P_i would be a simple Gaussian density with:

$$\zeta | P^*, P \sim \mathrm{N}(\hat{\zeta}, (\tilde{Z}'\tilde{Z})^{-1}),$$

with $\hat{\zeta} = (\tilde{Z}'\tilde{Z})^{-1})\tilde{Z}P^*$. The posterior distribution of P_i^*, conditionally on ζ and P_i is a truncated normal with:

$$P_i^* \sim \mathrm{TN}(\tilde{z}_i'\zeta, 1),$$

which is truncated at zero by the left if $P_i = 1$ or by the right if $P_i = 0$. A Gibbs sampler is devised by simulating alternatively P_i^* and ζ. A Maximum likelihood estimator serves at initializing the chain.[7] Inference for the decision to participate is done on all the observations (both positive and zero incomes). We then compute for positive wages the inverse of the Mills ratio (IMR). Therefore, we obtain m draws of the IMR which can be evaluated at the posterior expectation of ζ, so as to obtain $\lambda(\tilde{z}_i\bar{\zeta})$, and then inference on the wage equation is then conducted conditionally on the mean value $\lambda(\tilde{z}_i\bar{\zeta})$. This approach does not take into account the full uncertainty resulting from inference of the probit model. Therefore, we can directly include these m posterior draws of the IMR in the inference process for the wage equation.

5.2 The heteroskedastic model

After truncation, we have n positive observations for the wage noted y and X is the matrix of n observations and k covariates including x_i, the IMR and a constant term. The complete

[7] To draw a random number π from a truncated normal distribution of mean μ and unit standard deviation between bounds $a < \pi < b$, we apply the inverse sampling method:

1. Compute $\bar{a} = a - \mu$ and $\bar{b} = b - \mu$.
2. Generate r^{th} random number ξ^r from the uniform distribution.
3. Define $\bar{\xi}^r = (1 - \xi^r)\Phi(\bar{a}) + \xi^r\Phi(\bar{b})$.
4. Obtain $\pi = \Phi^{-1}(\bar{\xi}^r) + \mu$ which lies between a and b.

regression model is noted:

$$\log y = X\beta + v, \qquad v \sim N(0, \sigma^2 H).$$

The $n \times n$ matrix H represents the variance-covariance matrix of the error term v, this is a diagonal matrix:

$$H(\delta) = \text{diag}(h(z_1'\delta), \ldots, h(z_n'\delta)).$$

Let us recall that z_i contains the exogeneous variables x_i of the wage equation and $\tilde{z}_i$ the explanatory variables of the selection equation plus $\lambda(\tilde{z}_i'\bar{\zeta})$ and $\lambda(\tilde{z}_i'\bar{\zeta})^2$. Because we explicitly model heteroskedasticity, the posterior standard error of β will be correctly evaluated, contrary to the usual two-step regression of Heckman which requires a specific correction as detailed in Cameron and Trivedi (2005, page 550) for instance.

The likelihood function of y given X and the parameters is:

$$\begin{aligned} L(y; \beta, \sigma^2, \delta) &= \left(\prod_{i=1}^{n} (y_i)^{-1/2}\right) (2\pi)^{-n/2} \sigma^{-n} |H(\delta)|^{-1/2} \\ &\quad \times \exp -\frac{1}{2\sigma^2} (\log(y) - X'\beta)' H^{-1} (\log(y) - X'\beta). \end{aligned}$$

This likelihood function is identical to the one considered for instance in Bauwens et al. (1999, Chap. 7), Griffiths (2001), and Koop (2003, Chap. 6), except for the Jacobian of the transform of y into $\log y$. We select a non-informative prior on all the parameters as done in Griffiths (2001):

$$\pi(\beta, \sigma^2, \delta) \propto 1/\sigma^2,$$

so that the posterior density of the parameters is proportional to:

$$\begin{aligned} \pi(\beta, \sigma^2, \delta | y) &\propto \sigma^{-(n+1)} |H(\delta)|^{-1/2} \\ &\quad \exp -\frac{1}{2\sigma^2} \left[s_*(\delta) + (\beta - \beta_*(\delta))' M_*(\delta) (\beta - \beta_*(\delta)) \right], \end{aligned} \tag{19}$$

with:

$$M_*(\delta) = X' H^{-1}(\delta) X, \tag{20}$$

$$\beta_*(\delta) = M_*^{-1}(\delta) X' H^{-1}(\delta) \log(y), \tag{21}$$

$$s_*(\delta) = (\log(y) - X\beta_*(\delta))' H^{-1}(\delta) (\log(y) - X\beta_*(\delta)). \tag{22}$$

There are several ways of treating this posterior density in β, σ, δ (conditionally on the IMR computed at the mean value of ζ). Koop (2003, Chap. 6) derives each conditional density and proposes a Gibbs sampler with a Metropolis step for the conditional distribution of δ. Bauwens et al. (1999, Chap. 7) and Griffiths (2001) prefer to note that conditionally on δ, we recover the conditional posterior distributions of β and σ^2 which are an inverted gamma 2 and a Student:[8]

$$\pi(\sigma^2 | \delta, y) = f_{i\gamma}(\sigma^2 | n, s_*^2(\delta)) \tag{23}$$

$$\pi(\beta | \delta, y) = f_t(\beta | \beta_*(\delta), M_*(\delta), s_*^2(\delta), n), \tag{24}$$

[8]Notations for these two densities are provided in the appendix of Bauwens et al. (1999), together with procedures to draw random numbers from them.

while we can derive the marginal posterior density of δ by an analytical integration of (19) in β and σ^2. The result is immediately deduced as the constant of integration of the above Student times a term which comes from the likelihood function, so that:

$$\pi(\delta|y) \propto |H(\delta)|^{-1/2} s_*(\delta)^{-(n-k)/2} |M_*(\delta)|^{-1/2}.$$

This density does not belong to a known family. If the dimension of z_i is one and consequently also the dimension of δ, we can draw random numbers from $\pi(\delta|y)$ using a Metropolis algorithm and a Gaussian proposal. The mean and the variance of the proposal can be approximated by the posterior mode of δ and minus the inverse of the second order derivative of the log posterior density at this point, respectively. Note that this proposal is calibrated at a value which depends on the IMR evaluated at the mean value of ζ.

The independent Metropolis algorithm should be preferred to the random walk Metropolis used in Griffiths (2001) because it implies a much lower rejection rate as underlined in Bauwens et al. (1999, Chap. 3). Let us denote $\ell(\delta)$ the proposal and $\pi(\delta|y)$ the posterior density, the independent Metropolis algorithm can be implemented as follows:

1. Generate a proposal $\delta^p \sim \ell(\delta)$.
2. Compute the probability of acceptance as $p_a = \min\left(\frac{\pi(\delta^p|y)}{\pi(\delta^{(j-1)}|y)} \frac{\ell(\delta^{(j-1)})}{\ell(\delta^p)}, 1\right)$.
3. Generate a uniform random number u.
4. If $u \leq p$, then accept $\delta^{(j)} = \delta^p$, otherwise keep $\delta^{(j)} = \delta^{(j-1)}$.

Once we have drawn random numbers from $\pi(\delta|y)$, it is easy to generate random numbers for β and σ. We have simply to replace $M_*(\delta)$, $\beta_*(\delta)$ and $s_*(\delta)$ by $M_*(\delta^{(j)})$, $\beta_*(\delta^{(j)})$ and $s_*(\delta^{(j)})$ and then use the following densities for drawing values for σ^2 and β:

$$\pi(\sigma^2|\delta^{(j)}, y) = f_{i\gamma}(\sigma^2|n, s_*^2(\delta^{(j)})), \tag{25}$$
$$\pi(\beta|\sigma^2, \delta^{(j)}, y) = f_N(\beta|\beta_*(\delta^{(j)}), \sigma^2 M_*^{-1}(\delta^{(j)})). \tag{26}$$

This approach by direct sampling avoids, at least partly, the dependence in the draws which is inherent to MCMC.

The above method supposes that the IMR is fixed, precisely the IMR is supposed to be a vector of length n. In doing so, we have evaluated the IMR at the posterior mean of ζ. However, it turns out that this method ignores partly the uncertainty resulting from the first stage probit. Indeed, Bayesian inference for the probit model results in m draws of the IMR for each individual $i = 1, ..., n$. Consequently, conducting inference for the heteroskedastic model is more demanding as the set of regressors is not fixed in both the wage equation and the skedastic equation. They include respectively $\lambda(\tilde{z}_i'\zeta^{(j)})$ for x_i and $[\lambda(\tilde{z}_i'\zeta^{(j)}),\ \lambda(\tilde{z}_i'\zeta^{(j)})^2]$ for z_i where $\zeta^{(j)}$ is the j^{th} draw from the probit MCMC output. The delicate question is that we should change the proposal of the Metropolis step for each draw $\zeta^{(j)}$ in order to get a more efficient proposal. However, the cost involved might be greater than the gain in efficiency, so we prefer to keep the same proposal, obtained for the average ζ.

5.3 Bayesian testing

Once we have m draws for μ, σ and δ, we can easily transform these draws into draws of $g_t(p)$ and $G_t(p)$, where a grid of np values has been chosen for p. These draws are stored in two matrices with m rows and np columns. The posterior mean curve is obtained by taking the mean of those np columns while taking the 0.05 and 0.95 quantiles over each column provides 90% confidence bounds. This method is straightforward in a Bayesian context contrarily to the classical approach, as shown in Araar et al. (2009).

Once we have m posterior draws of the growth incidence and poverty growth curves, many hypotheses can be tested:

1. whether a distributional change has been welfare improving in terms of first-order (second-order) stochastic dominance i.e. $g_t(p) > 0$ ($G_t(p) > 0$) for all p,
2. whether a distributional change has been relatively pro-poor i.e. $g_t(p) > \gamma$ or $G_t(p) > \gamma$ for all $p \leq F(z)$,
3. whether every quantile has benefited equally from growth i.e. $g_t(p) - \gamma = 0$ for all p,
4. and finally, whether the distributional change of a group is preferred to that of the other group i.e. $g^T(p) > g^C(p)$ or $G^T(p) > G^C(p)$ for all p.

In this attempt, we only have to compute for each value p of the grid, the probability that the curve is greater than zero or γ. Remark that the average growth rate γ is computed using equation (15) where x is replaced by $\bar{x}$. These probabilities are simple to evaluate within a Bayesian framework. Let us take the example of the sign of $g_t(p) - \gamma$. A j^{th} Monte Carlo draw for this event is noted $\mathbb{I}(g_t^{(j)}(p) > \gamma_t^{(j)})$, where $\mathbb{I}(.)$ is the indicator function. We have simply to evaluate the sampling mean of:

$$\Pr(g_t(p) > \gamma_t) \simeq \frac{1}{m} \sum_{j=1}^{m} \mathbb{I}(g_t^{(j)}(p) > \gamma_t^{(j)}).$$

Remark that because γ appears also in the definition of $g_t(p)$ and $G_t(p)$, we have to use different draws when evaluating $g_t^{(j)}(p)$ and $\gamma_t^{(j)}$.[9] The appeal of parametric curves stems from the ease in which they can be interpreted. Particularly, the growth incidence curve is horizontal if the variances of both periods are equal i.e. $\sigma_{t-1} = \sigma_t$ in the simple log-normal model or $\sigma_t\sqrt{h(z_i'\delta_t)} = \sigma_{t-1}\sqrt{h(z_i'\delta_{t-1})}$ in the model with heterogeneity. Therefore, it is possible to assess welfare criteria directly from the posterior draws of the parameters.

6 Two stories of inequality dynamics

We want to measure the effect of having been exposed to Western-German broadcasts before the reunification on the evolution of inequality of former East-German citizens after the reunification.

[9] A simple reshuffling is sufficient.

The sample is composed of two exclusive groups, those who received Western television broadcasts while living in East-Germany in 1990 (treatment group) and those who did not received Western television signals while living in East-Germany in 1990. Berlin has been excluded from the sample.

6.1 Labour participation parameter estimates

We have seen in Table 2 that there was a significant proportion of zero labour incomes. It reflects the decision to participate onto the labour market. In order to correct for such potential selection bias, we first estimate the participation equation. Specifically, we have selected the following explanatory variables for explaining the decision to participate to the labour market: the age, the age squared, the gender (a dummy indicating one for women and zero for men), the number of years of education, the number of children in the household, public and social transfers received by the household and the market income from other household members. Remark that we allow the effect of gender to differ with respect to the number of children in the household. Some values for the variable education are missing, then we impute these values by taking the mean conditionally on the other exogenous characteristics. Table 4 provides the posterior mean and standard deviation of the labour participation parameters for 1992 and 1995. We get 10 000 draws of the parameters and 1 000 are used for warming the chain. Monetary quantities are scaled by the SOEP consumer price index and are thus expressed in thousands real euros.

	1992		1995	
	Mean	Std. dev.	Mean	Std. dev.
Constant	-2.113	0.260	-2.745	0.328
Age	0.236	0.013	0.203	0.015
Age^2	-0.305	0.016	-0.258	0.018
Gender	-0.600	0.061	-0.524	0.063
Education	0.047	0.011	0.113	0.012
No. children	0.127	0.039	0.207	0.043
Gender $\times$ No. child.	-0.221	0.049	-0.432	0.053
Public transfers	-0.114	0.006	-0.115	0.005
Social transfers	-0.105	0.006	-0.094	0.005
HH other income	-0.010	0.002	-0.009	0.002

Table 4: Inference for the labour participation equation

Several features are worth mentioning. Even if the signs of the posterior means do not change, there are some differences between the two periods in the magnitude of the coefficients and consequently in the distribution of the Mills ratios. As expected, gender and age are the two main determinants of labour participation even if education, the number of children in the household and the monetary resources of the households are found to be significant determinants of the participation too. Particularly, we found that the likelihood of participating decreases with age, the other sources of income and when being a female whereas it increases with the level of education. Remark that the number of children in the household increases the participation of males whereas it reduces that of women. We investigate also whether the treatment has affected the decision to work and we found no evidence of a significant effect. Remark that the wage equation is identified as long as some of these variables are not determinants of the wage (i.e. exclusion condition). Clearly, the public and social transfers received by the household and the market income from other household members should not affect one individual's wage while affecting negatively and significantly the probability of labour participation (i.e. relevant condition).

6.2 Wage equation parameter estimates

We use a conventional Mincer equation to explain the labour earnings including the level of education, the gender and the age (as a proxy for experience) to explain the natural logarithm of the wage. We differ from the standard model by including a potential treatment effect and correcting for self-selection by including the inverse Mills ratio (IMR). As discussed before, the latter requires to correct the variance by allowing for potential heteroskedasticity. In both, the wage equation and the skedastic equation we include the treatment variable. In doing so, we allow the treatment to affect both the mean labour income and its variance (as inequality relies only on the σ parameter in the log-normal model). The trajectory of both treated and controlled households is measured between 1992 (the first data for which we have income data) and 1995 (three years later). We do not consider for the while a later period, because as time elapsed, the control group experiences the effect of Western television and many other effects could have inferred into the changes in the income distribution of the two groups. Table 5 provides the posterior mean and standard deviation of the wage equation parameters for 1992 and 1995 while Table 6 provides the posterior moments for the skedastic equation parameters. We get 10 000 draws of the parameters and 1 000 are used for warming the chain. We get acceptance rates of 0.93 and 0.80 for the Metropolis step in the first and the second period, respectively. As before, monetary quantities are scaled by the consumer prince index and are expressed in thousands real euros. Remark that we only provide results for variables which are found to be significant.

	1992		1995	
	Mean	Std. dev.	Mean	Std. dev.
μ	7.539	0.185	7.890	0.279
Treatment	-0.002	0.025	-0.035	0.032
Age	0.069	0.009	0.060	0.013
Age^2	-0.076	0.011	-0.062	0.015
Gender	-0.243	0.024	-0.115	0.033
Education	0.069	0.005	0.062	0.006
IMR	-0.848	0.104	-1.142	0.116

Table 5: Inference on wage equation parameters

The average income growth rate over the period 1992-1995 resulting from the estimated model is measured as being 16.4% while the median growth rate is lower with 10.1%, showing

	1992		1995	
	Mean	Std. dev.	Mean	Std. dev.
σ^2	2.367	1.003	3.405	1.734
Treatment	-0.056	0.068	0.207	0.077
Age	-0.103	0.020	-0.139	0.022
Age^2	0.108	0.025	0.146	0.027
Education	-0.009	0.013	0.042	0.014
Pub. trans	0.081	0.010	0.071	0.012
IMR	1.920	0.297	1.525	0.211
IMR^2	-0.900	0.182	-0.316	0.073

Table 6: Inference on skedastic equation parameters

intuitively that growth was more distributed toward the rich rather than the poor.[10]

Several features are worth mentioning. The Mills ratio is significant and has a strong correcting effect confirming a non-negligible selection bias. The exposure to Western Television has no impact on the mean wage during both periods. However, the exposure to Western television has affected strongly inequality (the variance of the log wage) between 1992 and 1995, passing from $-0.056 \times 0.146 = -0.008$ to $0.207 \times 0.152 = 0.031$ (computed at the mean value of the skedastic function). It turns out that the part of the variable component of the variance has much increased. More generally, we found that age, gender and education affect significantly the mean wage while age, education and public transfers affect significantly the variance.

6.3 Estimated growth incidence curves

Once we have obtained m posterior draws of the parameters, we can draw m parametric growth incidence curves. Figure 2 displays the estimated growth incidence curve for the whole sample (both control and treatment groups) showing the growth of wages over the period for each quantile.

Because this curve is increasing over the quantiles, growth has been inequality increasing between 1992 and 1995. Furthermore, since the wage growth rate is negative for the first decile, this indicates that the situation of the poor has been worsened whereas the rich experienced the most important increase. This last assertion should be nuanced as the non-parametric estimate[11] does not display such an improvement for the last decile but rather a constant wage growth rate above the median. This discrepancy between both curves might not indicate a misspecification but instead the potential lack of reliability of the non-parametric growth incidence curve in the case of data trimming. As far as we know, the non-parametric growth incidence curve is sensitive to the number of observations and can behave erratically even with relatively important sample

[10] The mean and median are obtained by applying formulae of the log-normal given in (15) for the mean and being $\exp(\mu + x_i'\beta)$ for the median, computed for each MCMC draw with explanatory variables taken at their mean values.

[11] The parameter-free estimator of the GI curve is provided by the difference of the log of the empirical quantiles computed on a grid of 20 points between 0.05 and 0.95.

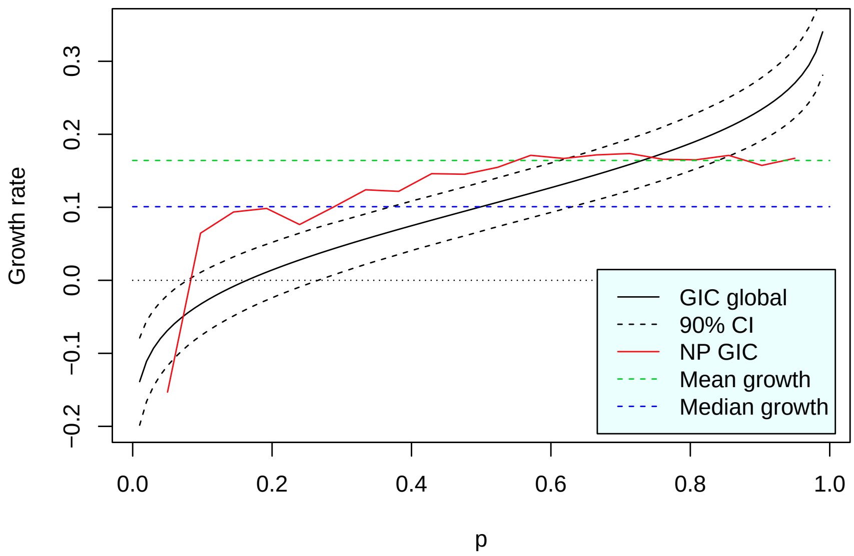


Figure 2: Growth incidence curve for average covariates (1992-1995)

sizes. This drawback amplifies as we are interested in the tails of the distributions. Therefore, we argue that the use of parametric growth incidence curve is necessary for small and medium sample sizes.

Table 7 reports the probability that the GIC is greater than a given threshold for each decile. This table confirms that the last decile benefited more of growth than the other deciles, and conversely, that the lower deciles benefited less of growth than the rich. In addition, Table 7 confirms that the wage of the first decile has certainly been reduced. Consequently, there is no first-order stochastic dominance of the second period wage distribution over that of first period.

	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
GIC $> \gamma$	0.00	0.00	0.00	0.00	0.01	0.10	0.36	0.74	0.97
GIC > 0	0.11	0.72	0.98	1.00	1.00	1.00	1.00	1.00	1.00

Table 7: Probability that the GIC is above a given threshold

6.4 Estimated treatment effect

Let us now display in Figure 3 the treatment and control growth incidence curves, showing how inequality have evolved in each group over the period. In other words, is there a significant difference in inequality evolution between the two groups? In this attempt, we have computed the treatment and control curves at the group characteristics means.

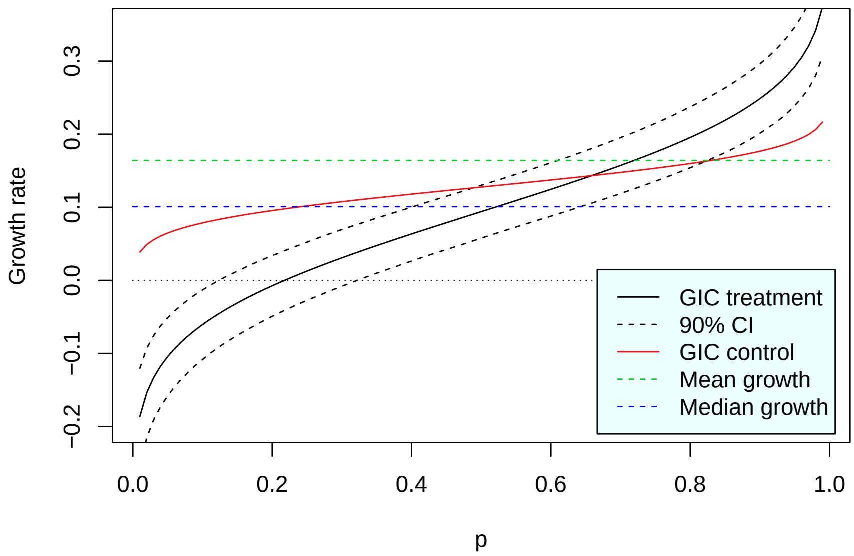


Figure 3: Growth incidence curves for treatment and control groups (1992-1995)

Clearly, the two groups have experienced substantial differences in inequality dynamics after the reunification. Indeed, the slope of the GIC curve of the treated group appears to be fairly steep, while that of the control group is much more horizontal. More precisely, even if the growth incidence curve of the control group is slightly increasing, the difference between the wage growth rate of the poor compared to that rich is small while being positive for both. This means that every quantile has benefited from growth (even if the poor has benefited less than the rich). This is a situation of first-order stochastic dominance wherein the wage distribution of the second period dominates that of the first period. This contrasts with the situation of the treatment group for which the situation of the poor has been worsened while that of the rich has been increased substantially. We have thus a complementary picture to the one given in Table 3 and also a much more contrasted one. In addition, we report in Table 8 the probability that a group GIC is greater than a given threshold for each decile and we also compare both curves each other. Formal statistical tests confirm that inequality in the treatment group has exploded and contrast sharply with the control group.

6.5 Counterfactual curves

Figure 3 compares the evolution of inequality between the control and the treatment groups. In doing so, we computed the curves at the respective group characteristics means. Thus, the difference between the two curves is due to either the effect of exposure to Western television or to differences in group characteristics means. In order to assess the causal effect of exposure to Western television, we compute what would be the growth incidence curve of the treatment group

Deciles	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
Treatment group									
Treatment $> \gamma$	0.00	0.00	0.00	0.00	0.01	0.09	0.39	0.80	0.99
Treatment $>$ median	0.00	0.00	0.01	0.10	0.39	0.77	0.96	1.00	1.00
Treatment > 0	0.02	0.38	0.90	1.00	1.00	1.00	1.00	1.00	1.00
Control group									
Control $> \gamma$	0.06	0.08	0.11	0.14	0.19	0.26	0.35	0.46	0.58
Control $>$ median	0.35	0.46	0.56	0.65	0.73	0.80	0.85	0.88	0.92
Control > 0	0.93	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Control $>$ Treatment	0.99	0.98	0.95	0.89	0.79	0.62	0.43	0.25	0.12

Table 8: Probabilistic comparison of treatment and control GIC

in the absence of the treatment (the counterfactual curve). Figure 4 provides the counterfactual

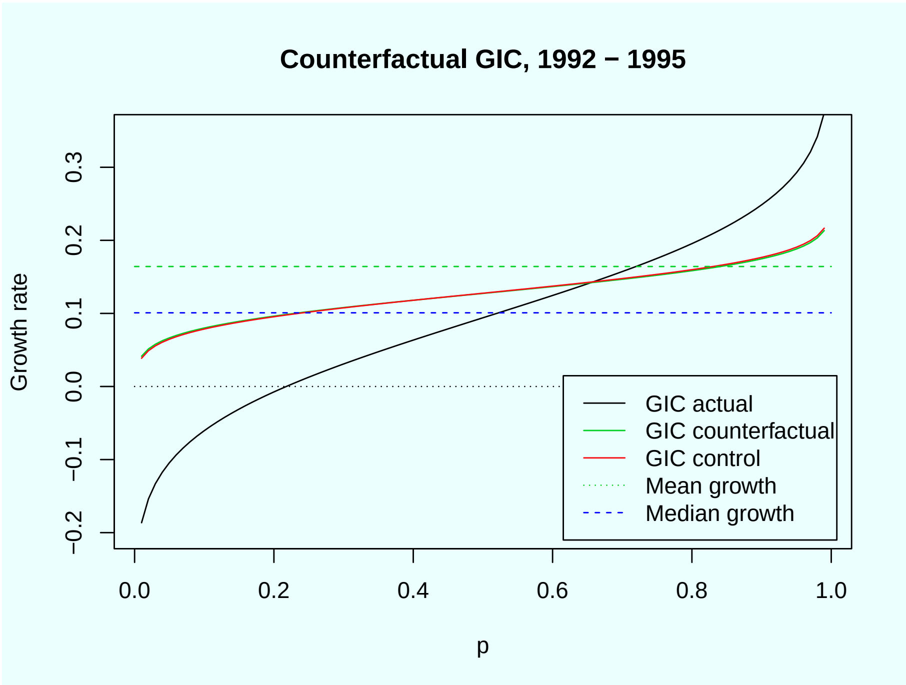


Figure 4: Counterfactual growth incidence curve (1992-1995)

growth incidence curve together with the actual curves of the treatment group and that of the control group. It turns out that in the absence of television exposure, the treatment group would have exactly the same curve than the control group. This means that all the difference between the treatment and the control group inequality dynamics is attributable to the effect of Western television exposure; the differences in group characteristics being negligible as shown in Table 9. In addition, we test whether the population mean group characteristics are equal using an Hotelling test. There is no statistical difference between the mean group characteristics in the first period, but we reject with confidence this assumption in the second period.

	1992		1995	
	Treatment	Control	Treatment	Control
Age	39.35	39.98	41.41	41.94
Education	11.98	12.23	12.24	12.59
Female	0.47	0.48	0.46	0.47
No. children	0.85	0.95	0.78	0.87
Public transfers	2.64	2.94	2.44	2.51
Social transfers	0.93	0.79	1.08	1.05
HH other income	18.92	20.04	19.41	21.99
IMR	0.173	0.167	0.242	0.232

Table 9: Mean group characteristics

6.6 Second-order stochastic dominance

Remains the question of whether the distributional changes due to television exposure has been favourable for individuals. Under the veil of ignorance, stochastic dominance is used to determine whether a distribution is preferred to another. Particularly, if a distribution first-order dominates another one, then the first distribution is preferred for any non-decreasing social utility function. As presented before, first-order stochastic dominance of the first period over the second period is verified if and only if the growth incidence curve is positive for every quantile. It follows that the change in the wage distribution of the treatment (control) group is preferred to that of the control (treatment) group if and only if the growth incidence curve of the former group is always above that of the later. Figure 3 and Table 8 shows that the growth incidence curve of the control group dominates that of the treatment group up to the fourth decile and then both curves intersect. Thus, we cannot conclude unambiguously about first-order stochastic dominance, one has to consider stochastic dominance at the second order implying to impose more normative conditions. Particularly, if a distribution second-order dominates another one, then the first distribution is preferred for any non-decreasing and concave (risk-averse) social utility function. Consequently, second-order dominance is weaker than the first order stochastic dominance and is more likely to be satisfied. As discussed before, second-order dominance can be inferred from the poverty growth curves. This argument is developed in Son (2004) as an argument for adopting poverty growth curves. Specifically, second-order stochastic dominance of the first period over the second period is verified if and only if the poverty growth curve is positive for every quantile. Similarly, the change in the wage distribution of the treatment (control) group is preferred to that of the control (treatment) group if and only if the poverty growth curve of the former group is always above that of the later.

Figure 5 provides the poverty growth curves for the treatment and control groups. The poverty growth curve of the control group is always statistically greater or equal to that of the treated group as confirmed in Table 10. Therefore, we can conclude that distributional changes happening in the control group are preferred (for any risk-averse utility function) to the changes in the treated group. In other terms, Western television exposure has led to an unambiguous increase in poverty and inequality.

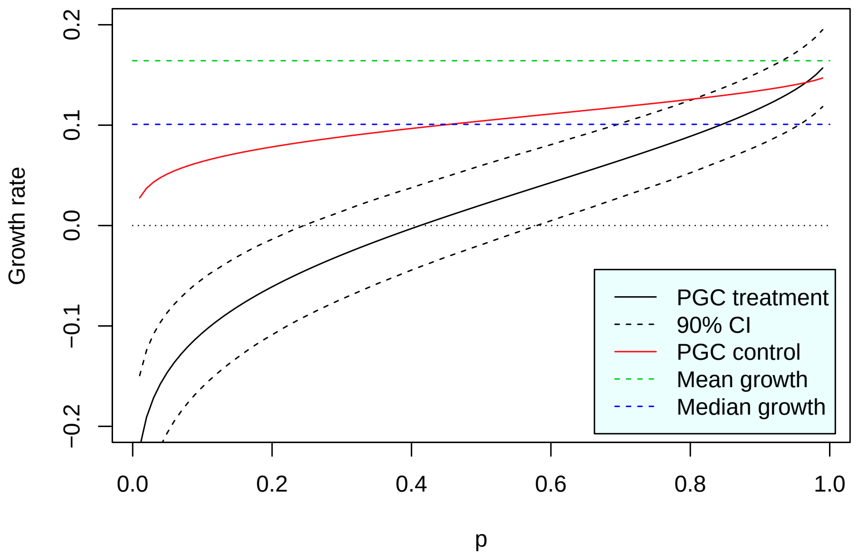


Figure 5: Poverty growth curves for treatment and control groups (1992-1995)

	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
$PGC_c(p) > PGC_t(p)$	0.99	0.99	0.99	0.98	0.96	0.94	0.89	0.81	0.67

Table 10: Probability of PGC dominance

6.7 Longer term effects

We demonstrated that Western television exposure led to an unambiguous increase in poverty and inequality between 1992 and 1995. However, a fundamental question is of course to know how long will last the effect of the treatment. After the German reunification, the control group was exposed to Western television and East-Germany experienced deep structural changes with the disappearance of old industries. Thus the effect of aspirations had been mixed with many other factors which have affected both the treated and the control groups. Therefore, we might expect that the effect of treatment has been diluted over time.

As the period of time considered goes on, the issue of sample attrition exacerbates. The respective sizes of the treatment and control groups are indicated in Table 11, from which we have excluded the zero labour incomes so that the resulting samples sizes for the control group are going to be very small limiting the reliability of long term inference.

Figure 6 provides the growth incidence curves for the period 1992-1995/98. More precisely, we use 1992 as initial period and we pool the waves 1996-1998 as final period for the sake of not entailing too much the statistical power. Clearly, the treatment effect disappears when we extend the final period after 1995 in such a way that both treatment and counterfactual curves are not statistically different. We also consider each year separately and the results remain qualitatively

Dates	1992	1995	1997	2000	2005
All	2698	2112	1822	1540	1093
Treatment	2120	1677	1451	1225	884
Control	578	435	371	315	209

Table 11: Sample attrition

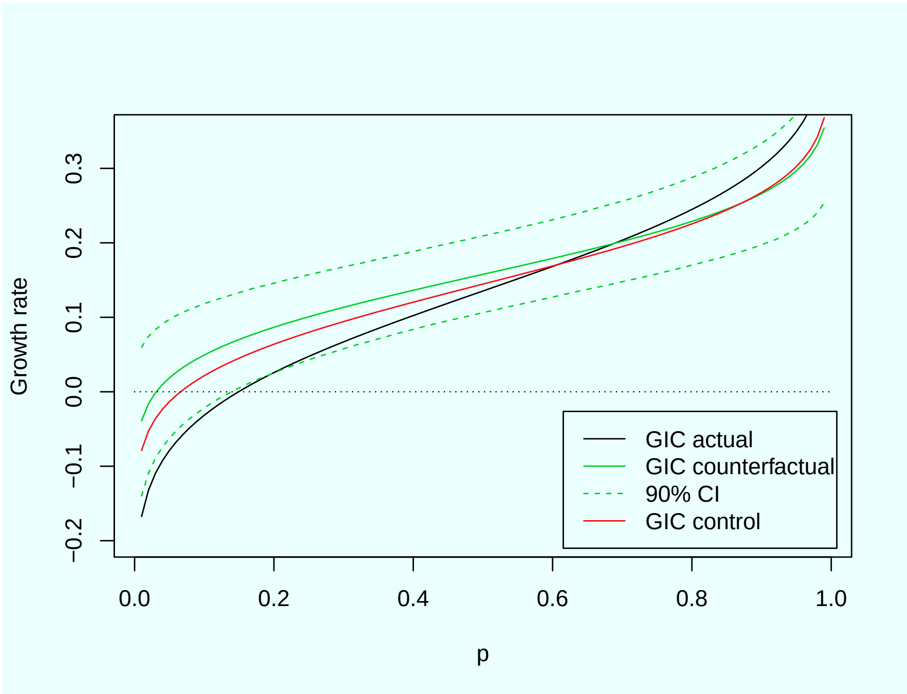


Figure 6: Long term growth incidence curves 1992-1995/98

similar even if the curves tend to behave erratically as the sample size reduces.

7 Conclusion

A growing interest has been devoted on how social aspirations could affect inequality. Indeed, social aspirations tend to influence individual decisions which, in turn, affect inequality and poverty. We empirically examine the effect of aspirations on inequality through an exogenous exposure to Western television channels in East-Germany before the reunification. In this attempt, we propose an heteroskedastic parametric model for income with a treatment effect while taking into account potential sample selection into the labour market. We derive analytical formulae for the growth incidence curve of Ravallion and Chen (2003) and the poverty growth curve of Son (2004) for the log-normal distribution. Based on those curves, we provide Bayesian inference and a set of tests related to stochastic dominance criteria. We find evidences that aspirations - through exposure to Western German broadcasts - have significantly affected inequality. We provide evidences that this effect was detrimental in terms of inequality and poverty. However, we cannot conclude about the persistence of this effect after 1995. Although, we assumed that the main mechanism through which Western television exposure affects inequality operates through social aspirations, we cannot definitively exclude that alternative channels (e.g. information) may have affected inequality as well.

The econometric model we use in this paper is not exempt of potential criticisms and shortcomings. Of course, our approach relies on the quality of the parametric assumption. We argue that the log-normal distribution is very convenient for modelling the income distribution and deriving analytical expressions for the growth incidence curve and the poverty growth curve. Although, we argue that our results are not driven by this constrained specification, the use of a semi-parametric mixture approach will make the distributional assumption much more flexible. Another point relies on the fact that we measure only changes in the income distribution while ignoring income mobility and individual lots. Even if this relaxes the data requirements as it is sufficient to use repeated cross-section data for estimating distributional changes, a longitudinal approach will allow to control for fixed unobserved individual effects. We let these issues for further researches.

References

Aitchison, J. and Brown, J. A. (1957). The lognormal distribution with special reference to its uses in economics. *The American Economic Review*, 48(4):690–692.

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679.

Anderson, G., Pittau, M. G., and Zelli, R. (2014). Poverty status probability: a new approach to measuring poverty and the progress of the poor. *The Journal of Economic Inequality*, 12(4):469–488.

Araar, A., Duclos, J.-Y., Audet, M., and Makdissi, P. (2009). Testing for pro-poorness of growth, with an application to Mexico. *Review of Income and Wealth*, 55(4):853–881.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263.

Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica*, 55(4):749–764.

Bauwens, L., Lubrano, M., and Richard, J.-F. (1999). *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, Oxford.

Boenisch, P. and Hyll, W. (2015). Television role models and fertility: Evidence from a natural experiment. SSRN Scholarly Paper ID 2611597, Social Science Research Network, Rochester, NY.

Bourguignon, F. (2011). Non-anonymous growth incidence curves, income mobility and social welfare dominance. *The Journal of Economic Inequality*, 9(4):605–627.

Bursztyn, L. and Cantoni, D. (2016). A Tear in the Iron Curtain: The Impact of Western Television on Consumption Behavior. *Review of Economics and Statistics*, 98(1):25–41.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics, Methods and Applications*. Cambridge University Press.

Chancel, L. and Piketty, T. (2017). Indian income inequality, 1922-2014: From British Raj to billionaire Raj? Working Paper 2017/11, WID.world.

Chong, A. and Ferrara, E. L. (2009). Television and Divorce: Evidence from Brazilian Novelas. *Journal of the European Economic Association*, 7(2-3):458–468.

Clark, A. E. and Senik, C. (2010). Who Compares to Whom? The Anatomy of Income Comparisons in Europe. *The Economic Journal*, 120(544):573–594.

Davidson, R. and Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, 68(6):1435–1464.

Duclos, J.-Y. (2009). What is Pro-Poor? *Social Choice and Welfare*, 32(1):37–58.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3):761–766.

Foster, J. and Shorrocks, A. (1988). Poverty orderings. *Econometrica*, 56(1):173–177.

Fourrier-Nicolai, E. and Lubrano, M. (2017). Bayesian inference for TIP curves: An application to child poverty in Germany. Working Paper 2017-10, AMSE, Marseille.

Frey, B. S., Benesch, C., and Stutzer, A. (2007). Does watching TV make us happy? *Journal of Economic Psychology*, 28(3):283 – 313.

Friehe, T., Muller, H., and Neumeier, F. (2017). The effect of Western TV on crime: Evidence form East Germany. Working Paper 10-2017, Joint Discussion Paper Series in Economics.

Gastwirth, J. (1971). A general definition of the Lorenz curve. *Econometrica*, 39(6):1037–1039.

Genicot, G. and Ray, D. (2017). Aspirations and Inequality. *Econometrica*, 85(2):489–519.

Griffiths, W. (2001). Heteroskedasticity. In Baltagi, B. H., editor, *A Companion to Theoretical Econometrics*, chapter 4, pages 82–100. Blackwell Publishing Ltd, Oxford.

Grimm, M. (2007). Removing the anonymity axiom in assessing pro-poor growth. *The Journal of Economic Inequality*, 5(2):179–197.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Hyll, W. and Schneider, L. (2013). The causal effect of watching TV on material aspirations: Evidence from the valley of the innocent. *Journal of Economic Behavior & Organization*, 86:37–51.

Jenkins, S. P. and Van Kerm, P. (2006). Trends in income inequality, pro-poor income growth, and income mobility. *Oxford Economic Papers*, 58(3):531–548.

Jensen, R. and Oster, E. (2009). The Power of TV: Cable Television and Women's Status in India. *The Quarterly Journal of Economics*, 124(3):1057–1094.

Knight, J. and Gunatilaka, R. (2012). Income, aspirations and the hedonic treadmill in a poor society. *Journal of Economic Behavior & Organization*, 82(1):67 – 81.

Koop, G. (2003). *Bayesian Econometrics*. Wiley, Chichester, England.

Lakner, C. and Milanovic, B. (2016). Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession. *World Bank Economic Review*, 30(2):203–232.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219.

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics*, 120(3):963–1002.

Marin, J.-M. and Robert, C. P. (2007). *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Springer Texts in Statistics. Springer.

Novokmet, F., Piketty, T., and Zucman, G. (2017). From Soviets to Oligarchs: Inequality and Property in Russia, 1905-2016. Working Paper 23712, National Bureau of Economic Research.

Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1):53–68.

Ravallion, M. and Chen, S. (2003). Measuring pro-poor growth. *Economics Letters*, 78:93–99.

Shorrocks, A. F. (1983). Ranking income distributions. *Economica*, 50(197):3–17.

Son, H. H. (2004). A note on pro-poor growth. *Economics Letters*, 82(3):307 – 314.

Stutzer, A. (2004). The role of income aspirations in individual happiness. *Journal of Economic Behavior & Organization*, 54(1):89 – 109.

Veblen, T. (1899). *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. The Macmillan Company.

Appendix

A Television and historical context

At the end of World War II, Germany was split in two zones. While the three Western occupation zones (American, British and French) united economically and politically to form the Federal Republic of Germany (FRG) in 1949, the Soviet zone took a separate path, resulting into the creation of the German Democratic Republic (GDR), a socialist economy resolutely linked with the Soviet Union. Thus, the borders between East and West-Germany arose from the position of the occupying forces at the end of the war. The Cold War and the Iron Curtain have anchored the FRG in the Western Bloc (the United-States and its allies) and the GDR in the Eastern Bloc (the Soviet Union and its allies).

While the United States invested colossal means via the Marshall plan to rebuild war-torn regions (mainly the United Kingdom, France and West-Germany) but also to prevent the spread of Communism, the Soviet Union refused the plan benefits, blocked the benefits to Eastern Bloc countries while proposing the Molotov plan. These diplomatic manoeuvres have been amplified by ideological confrontations which have deeply affected East and West-Germany. While individualism and liberalism were promoted in the FRG by the Marshall plan, collectivism and totalitarianism ruled the East-German society. Individual satisfaction and private freedom were the main concern in the West whereas a centrally planned economy was in force in the East where individual aspirations were neglected.

In this context of political tensions between the East and the West, there was a tear in the Iron Curtain through the medium of television as presented in Hyll and Schneider (2013) and Bursztyn and Cantoni (2016) for instance. While beginning their broadcasts in the same year, 1952, the West and East public TV networks, ARD and DFF respectively, took very different paths following the dynamics of the Cold War. While the West public TV was founded by radios of the three Western occupation zones as a public federal institution promoting diversity of public opinion and the renunciation of state influence, its East counterpart was used as a mean for indoctrination. Indeed, East-German TV had to comply with the political authority and faced a restrictive censorship in order to give support for policies in force. Unlike Western TV, Eastern broadcasts promoted fertility and labour participation while reproving attachment to consumption goods, deviant and violent behaviours (e.g. Boenisch and Hyll 2015, Friehe et al. 2017, Bursztyn and Cantoni 2016). In this context of censorship, Western broadcasts were an important source of information for GDR citizens. Indeed, several West-German television broadcasting transmitters were placed next to the East-German border in order to maximize availability of Western broadcasts in East-Germany. The better perceived quality of Western news and entertainment programs have been pushed forward as the main determinants of such preference for Western broadcasts.

Reunification and the collapse of the GDR was as unexpected as quick. It started with the historic fall of the Berlin Wall in November 1989 and by October 1990 the economic and political systems were unified. More precisely, East-Germany became part of the FRG and the economic and political system of the West was transferred to the East. In December 1990, the West-German public TV network took over the former East-German public TV. At the time of its fall, the GDR was the most developed economy of the Eastern Bloc but was nonetheless decrepit by Western standards, with a barely competitive industrial structure, severe deficiencies in the production and distribution of goods, burdened with a high level of external debt required to keep the living standards of East-Germans high. Nowadays, socio-economic differences still persist between former West and East parts of Germany.