Which Tail Matters? Inequality and Growth in Brazil[*]

Stephan Litschig
Graduate Institute for Policy Studies

María Lombardi
University of Gothenburg

September 2017

Abstract

We estimate the effect of initial income inequality on subsequent income per capita growth using sub-national data from Brazil over the period 1970-2000. Holding initial income per capita and standard confounders constant, we find that places with higher initial inequality exhibit higher subsequent growth. This effect is entirely driven by the lower tail of the initial income distribution: compared to more equal places, sub-national units with a higher share of income going to the middle quintile at the expense of the bottom quintile grow more rapidly, while places with a higher share of income going to the top quintile at the expense of the middle quintile get no growth boost at all. We document that both physical and human capital accumulation in places with higher inequality in the lower tail of the initial income distribution outpace capital accumulation in more equal places, while inequality in the upper tail of the distribution is uncorrelated with subsequent physical or human capital growth. These results are consistent with theories on credit constraints and setup costs for human and physical capital investments.

JEL Classifications: D3, O1, O4.
Keywords: Income inequality, economic growth.

1   Introduction

A series of seminal theory papers propose different channels through which a society's degree of initial economic inequality might impact subsequent income per capita growth. These channels include aggregate savings and investment (Bourguignon, 1981), human and physical capital accumulation (Galor and Zeira, 1993; Banerjee and Newman, 1993; Aghion and Bolton, 1997; Barro, 2000; Galor and Moav, 2004), and income redistribution and social unrest (Persson and Tabellini, 1994; Alesina and Rodrik, 1994; Benabou, 1996; Esteban and Ray, 2000; Campante and Ferreira, 2007). Existing cross- and within-country studies on the relationship between inequality and growth have produced effect estimates ranging from negative to zero and positive as further discussed below. While research design and data limitations may account for some of this variability, it is also possible that the effect of inequality on growth is genuinely heterogeneous. Indeed some of the mechanisms above have different implications for the effect of income inequality on growth, depending on whether the middle class is richer at the expense of the poor or the rich are richer at the expense of the middle class. Yet empirical work has almost exclusively focused on overall inequality.

This paper investigates whether inequality originating from the lower as opposed to the upper tail of the income distribution has different effects on subsequent income per capita growth. Greater inequality as measured by commonly used metrics (e.g. the Gini coefficient) can result from higher dispersion in different parts of the income distribution, as illustrated in Figure 1. In Panel A, a theoretical redistribution of income from the bottom to the middle quintile (i.e., the transition from the Lorenz curve displayed in the solid line to that of the dashed line) implies higher overall income inequality as captured by the Gini coefficient. However, the exact same increase in overall inequality can be achieved by redistributing a portion of total income from the middle to the top quintile, as shown in Panel B. Most existing empirical work has effectively treated the variation in overall inequality the same, irrespective of the tail it originates from, despite the fact that the theory suggests that impacts on subsequent growth may differ. For example, in models with credit constraints and setup costs for human (Galor and Zeira, 1993) and physical capital investments (Barro, 2000), it is conceivable that only inequality in the lower tail matters for growth.

Consider a stylized economy with three groups of equal size (the poor, the middle class, and the rich) and the same income within each group. Now assume that the incomes of the poor and the middle class are initially too low to overcome the setup costs for investing in either human or physical capital. Put differently, both the poor and the middle class cannot borrow enough to make the relatively large investments that would be required to make a profit. Now consider another economy with the same income per capita but with higher inequality at the bottom, i.e. the middle class is richer while the poor are poorer. In this second economy, the middle class might be rich enough to overcome the setup costs and make profitable investments in human and physical capital, thus making the second economy richer than the first economy in the long run. Finally consider a third economy, again with identical income per capita but higher inequality at the top, i.e. the rich are richer at the expense of the middle class. Since human and physical capital investments are as constrained as in the more equal first economy, growth will be similarly limited.

Using sub-national data from Brazil over the 1970-2000 period we first establish that holding initial income per capita and a host of standard confounders constant, places with higher initial income inequality as measured by the Gini coefficient exhibit higher subsequent income per capita growth. Most of the effect materializes by 1991, i.e. there is only a level effect, not permanently higher growth. We then propose a simple approach to distinguish between the growth effects of inequality originating from the bottom versus the top of the initial income distribution. The key idea is to include quintile income shares instead of the Gini coefficient in an otherwise standard cross-sectional growth regression, allowing for hypothetical income redistributions from the two tails towards the (omitted) middle quintile while holding other income shares and mean income constant. We find that the positive effect of overall inequality on subsequent growth is entirely driven by inequality in the lower tail of the income distribution: compared to more equal places, sub-national units with a 3 percentage point (one standard deviation) higher share of 1970 income going to the middle quintile at the expense of the bottom quintile experience about 3 percent higher income per capita by 2000. In contrast, places with a higher share of income going to the top quintile at the expense of the middle quintile get no growth boost at all compared to more equal places.

The differential effects of bottom versus top inequality are remarkably consistent with our evidence on human and physical capital accumulation. We find that places with a higher share of income held by the middle quintile at the expense of the bottom quintile experience higher subsequent growth in the real value of capital stocks held by businesses across all sectors of the economy. On the other hand, a higher share of income held by the top quintile at the expense of the middle quintile is not associated with increased physical capital accumulation. Moreover, human capital accumulation in places with higher inequality in the lower tail of the initial income distribution also outpaces places where the bottom quintile is richer, while inequality at the top of the distribution is uncorrelated with subsequent human capital growth. Other channels might also be at work. For example, higher inequality at the bottom might increase aggregate savings, and, in partly closed economies, aggregate investment. Higher inequality at the top on the other hand would affect aggregate savings only little if at all if the propensity to save is decreasing in income. Similarly, redistributive policies carried out at the local level may respond differently to lower- vs. upper-tail inequality. Unfortunately, we lack data on local savings or redistributive policies in the early 1970s.

Our paper builds on an extensive empirical literature linking overall income inequality and subsequent income per capita growth. Existing evidence is largely inconclusive and due to data limitations there is typically very little evidence on the mechanisms linking initial inequality to subsequent growth.[1] The most closely related study to ours is Voitchovsky (2005) which uses the 90/75 income percentile ratio as a proxy for inequality at the top of the income distribution, and the 50/10 income percentile ratio to proxy inequality at the bottom. For a sample of 21 developed countries, the study shows that under some specifications, inequality at the bottom is negatively correlated with growth, and inequality at the top has a positive correlation. The main conceptual difficulty

---

[1]Early cross-country studies had found a negative relationship between inequality and growth (Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996). However, this negative correlation has proven to be non-robust to the inclusion of additional explanatory variables, and the lack of comparability of inequality measures across countries potentially results in attenuation and other measurement error bias (Forbes, 2000). This last issue has been overcome by a series of studies exploiting variation in inequality across sub-national units (e.g., Partridge, 1997; Panizza, 2002; Benjamin et al., 2011). Employing a higher quality dataset drawn from household surveys, Deininger and Squire (1998) find a negative cross-country correlation between initial income inequality and growth, whereas Barro (2000) finds a positive correlation between the Gini coefficient and average income growth in rich countries, but a negative correlation in poorer countries. A second strand of studies using higher quality data as well as panel data estimations find a positive (Partridge, 1997; Li et al., 1998; Forbes, 2000), negative (Panizza, 2002; Benjamin et al., 2011) and negative though often insignificant (Voitchovsky, 2005; Ferreira et al., 2014) correlation between Gini and growth.

with the Voitchovsky (2005) study is that the regression specifications typically include percentile ratios along with the Gini coefficient in the same equation. But a higher 90/75 income percentile ratio while keeping the Gini coefficient constant necessarily implies that inequality must be lower in other parts of the distribution. As a result, it is not clear what the coefficient on the 90/75 income percentile ratio is picking up. A similar issue arises in Ravallion (2012), which explores the impact of various parameters of the initial income distribution on income per capita growth and poverty reduction in a large sample of developing countries. The regression specification sometimes includes the initial poverty rate along with the Gini coefficient in the same equation. But holding initial income per capita constant, countries with a higher poverty rate are also those with higher overall inequality, as discussed in that study. Moreover, holding both average income and overall inequality constant implies that countries with a higher poverty rate must have less inequality somewhere else in the distribution, which further complicates the interpretation of the coefficients.

The main contribution of our study is its conceptually straightforward approach to analyze the relationship between left and right tail inequality and subsequent outcomes. By replacing the Gini coefficient with the quintile income shares as our main regressors, we exploit variation in inequality originating from either tail while keeping initial average income and the other income shares constant. As illustrated in panels A and B of Figure 1, we exploit quantitatively identical differences in income inequality arising from opposing sides of the income distribution. This implies that the difference in effects of bottom versus top inequality we find is not driven by treating inequality in the two tails differently. And because our regressions hold income per capita constant, places with a lower share of income going to the poor and a higher share going to the middle class are places where the poor are poorer and the middle class is richer not only in relative but also in absolute terms. This is important because the credit market imperfections cum setup cost theory is based on absolute income levels. Another advantage of our setting is that we draw on homogeneously collected census data from a single country. Thus, unlike existing cross-country studies, we do not face a tradeoff between data quality and sample size, and our results are less prone to measurement error bias. Ours is also the first study to look at the effects of bottom versus top income inequal-

4

ity in a developing country context. An additional advantage of our setting is that by comparing sub-national units within the same country and state, we can abstract from differences in institutions at the federal and state level which might be correlated with initial inequality and income growth. Last but not least, our study also provides the first direct evidence on human and physical capital accumulation linking initial income distribution to subsequent economic growth.

A potential drawback compared to cross-country studies is that our results could be driven by migration, whereby places with high initial inequality at the bottom experience higher out-migration of the poor and thus higher income per capita among remaining residents in future periods, for example. It turns out, however, that the effect of initial income inequality on in- or out-migration flows is close to negligible in practice as further discussed below. We also show that our results are unlikely to be driven by differential measurement error at the bottom versus at the top of the initial income distribution. And as in any observational study there is the possibility that our results are driven by some unobserved confounder, such as heterogeneity in local tastes for equality for instance. We show, however, that our estimates are almost unchanged if we adjust them to account for potential selection on unobservables as proposed in Oster (Forthcoming). Together with our evidence on human and physical capital accumulation, our study thus provides reasonably well-identified estimates of the link between income inequality and growth.

The paper is organized as follows. Section 2 describes the Brazilian setting, and section 3 describes the data and presents summary statistics. Section 4 discusses our approach to analyze the relationship between left and right tail inequality and growth and how we deal with potential confounding factors. Section 5 presents and discusses our main results, and section 6 presents evidence on mechanisms. Section 7 presents the results of multiple robustness checks, and section 8 concludes with a discussion of external validity.

2   The Brazilian Setting

The starting point of our analysis is 1970, which is dictated by the availability of comparable income data over time as further discussed below. Our units of analysis are the 3,659 Brazilian *Áreas*

Mínimas Comparáveis (AMCs), which are roughly equivalent to Brazil's municipalities in 1970. On average, AMCs had about 25 thousand inhabitants at that time. And while today Brazil is a middle-income country with a large urban population, this was by no means true in 1970 when a large fraction of its population lived in poverty, and more than half resided in rural areas. The level of education was also extremely low. In particular, AMCs had an average educational attainment of individuals above 25 years old of only 1.37 years, and an iliteracy rate of 44% for people above 15.

Since 97% of individuals who worked or studied did so in their municipality of residence, several mechanisms driving the relationship between income inequality and growth should operate within AMCs. While there was ample room for growth driven by the accumulation of human and physical capital, opportunities for investment were rather limited for households at the bottom of the wealth distribution, since access to credit was not widespread. For example, in the agricultural sector where 42% of the workforce was employed in 1970, only 12% of establishments received credit during that year.[2] Credit constraints were therefore likely binding for a large fraction of Brazilians in our period of study. Together with setup costs, inequality at the lower or upper tail of the income distribution might therefore lead to very different growth dynamics as argued in Galor and Zeira (1993), for example.

An important part of the literature has devoted attention to the role of political forces in explaining the relationship between inequality and economic development (e.g., Persson and Tabellini, 1994; Alesina and Rodrik, 1994; Benabou, 1996, among others). In particular, these studies posit that more unequal societies face higher pressure for redistribution, which in turn generates distortions and hampers growth. Although Brazil was under a military dictatorship from 1964 until 1985, local elections were still held in most municipalities. Furthermore, while only 2.6% of total revenues were raised by municipal taxes, around 12%-17% of total public spending was done by municipal governments (Hagopian, 1996). So even though within-AMC inequality in the 1970s could not impact local taxation in a relevant way, it might still affect the composition of spending and thus economic development. While we do not mean to play down the role of redistributive policies in

---

[2]These figures are based on our own calculations using the 1970 population and agricultural censuses. Information on access to credit for households or firms in other sectors is not available for this period.

mediating the effect of local inequality on subsequent growth, it is not clear from a theoretical perspective how inequality generated at the bottom as opposed to the top of the income distribution would interact with spending decisions at the AMC level. Furthermore, lack of information on spending at the local level for this period does not allow us to explore this issue further.

3   Data and Descriptive Statistics

Our analysis relies on the 25% sample of the 1970 and 1980 Brazilian censuses obtained from the Brazilian Statistical Agency (*Instituto Brasileiro de Geografia e Estadística*, IBGE), and on AMC-level statistics published by IPEA (*Instituto de Pesquisa Econômica Aplicada*).[3] The starting point of our analysis is 1970 since this is when the first round of the Brazilian census with precise information on individual incomes was conducted.[4] Our units of analysis are the 3,659 Brazilian AMCs, which are themselves based on all existing municipalities from 1970 to 2000. Since many municipalities split or merged with others after 1970, doing our analysis at the AMC-level allows us to keep the borders constant and follow the same geographical units over time.[5]

When working with the 1970 25% census sample we first match the 3,974 municipalities appearing in this census to their corresponding AMCs.[6] This census investigated the monthly income for all individuals 10 years and older and asked for: (i) the income of the last month for those who earn a fixed income (e.g., salaries, pensions, etc.); (ii) the average monthly income in the last twelve months for those who receive variable income (e.g. professionals' fees, sale and brokerage commissions, payments for services rendered, etc.); and (iii) the monthly average of other regular sources of income such as routine donations, rents, dividends, etc. We construct the per capita family income distribution for each AMC in 1970 by dividing the sum of the individual incomes of all family members living in the same household by the number of family members.[7] This way,

---

[3] Available at http://www.ipeadata.gov.br/.

[4] In the previous census round in 1960, income was reported in only eight categories.

[5] Brazil had 3,974 municipalities in 1970, and 5,507 by 2000.

[6] We match municipality and AMC codes using the Data Zoom program developed by the Department of Economics at PUC-Rio, available at http://www.econ.puc-rio.br/datazoom/english/.

[7] Only 1.68% of individuals who report having a source of income do not report their earnings.

all family members living under the same roof have the same per capita income. We exclude from our analysis those individuals living in collective dwellings (e.g. hotels, hospitals, nursing homes), which amount to 1.89% of our sample. We also exclude individuals living in a private dwelling who are unrelated to the family head (tenants and domestic servants) and who account for 2.21% of all individuals. We then construct three main indicators from each AMC's per capita family income distribution, using the appropriate expansion weights provided by IBGE. First, we calculate the average per capita family income in 1970, which we express in R$ of 2000. Second, we construct the 1970 AMC Gini coefficient,[8] and third the share of total AMC income held by each of the quintiles. We also calculate an approximation to the Gini coefficient using these quintiles shares.[9] Unlike subsequent censuses, incomes above Cr$ 9,998 are top-coded at this value,[10] affecting 0.04% of employed individuals. As a robustness check, we adjust top-coded incomes, multiplying them by a factor of 2.15 so that individual incomes in the top 20% follow a Pareto distribution.[11] We also use the 1970 census to compute the share of occupied individuals working in each of the 16 economic sectors detailed in the census, which we use as controls in the robustness checks we perform in Section 7.1.[12]

We apply the same procedure to the microdata from the 1980 25% long-form sample in order to obtain the per capita family income distribution of each AMC, and in turn compute a series of per capita income percentiles and poverty rates for each AMC. For computing poverty rates, we use three different poverty lines: (i) half of the Brazilian minimum wage in September 1991; (ii) US$ 2 a day at 2005 PPP, which is the median poverty line amongst developing countries based on a compilation of national poverty lines in Ravallion et al. (2009); and (iii) US$ 1.25 a day at 2005 PPP, the mean poverty line for the poorest 15 countries. The first of these was obtained from IPEA, whereas the

---

[8]We use the *ineqdec0* code written by Stephen Jenkins for this calculation.

[9]Define Qn as the share of total AMC income held by quintile n. Then Gini $\approx 0.8 \times$ [Q5+0.5Q4−0.5Q2−Q1].

[10]All figures in the 1970 and 1980 census are reported in Cruzeiros (Cr$), Brazil's currency at the time. We converted all figures to Brazilian Reais (R$) of 2000 using the guidelines employed by the 1998 "Atlas de Desenvolvimento Humano no Brasil."

[11]This methodology is commonly used by researchers working with CPS data in the US. Examples include Katz and Murphy (1992), Autor et al. (2008), and Autor and Dorn (2013).

[12]These sectors are agriculture and forestry; gathering of wild growing products; hunting and fishing; mining and quarrying; manufacturing; construction; public utilities; wholesale and retail trade; services; transporting and communications; education, health and social activities; public administration, legislation and justice; national defense and public safety; real estate, financial and insurance activities; liberal professions; and other activities.

others were taken from Ravallion (2012). We also rely on the 1980 census 25% sample to study the migration patterns across AMCs between the 1970 and 1980 censuses. More specifically, we compute immigration and emigration rates for each AMC between 1970 and 1980. Since the 1980 census asks individuals how long they have been living in their current municipality, we count all individuals in a particular AMC who report that they were not living in their current municipality in 1970 as immigrants. Individuals who are younger than 10 years old in 1980 and belong to a family in which the head is an immigrant are considered immigrants as well. We calculate an AMC's immigration rate as the ratio between the number of immigrants in 1980 and the AMC's total population in 1970. Furthermore, the census also asks people who have been living in their current municipality for less than 10 years to specify the municipality in which they were previously residing. Thus, for each AMC, we can calculate the number of people who were living there in 1970 and left. We use this to calculate the emigration rate, which is simply the number of emigrants of an AMC divided by the 1970 population. A caveat for this measure is that the municipality of origin is missing for approximately 19% of all immigrants. Since we cannot trace these people to their municipality of origin, our emigration rate does not include these observations in the numerator. We also calculate AMCs' fertility and mortality rates, to uncover the population dynamics in this period. The fertility rate is the ratio between the number of AMC natives who are less than 10 years old in 1980 and the population in 1970. The mortality rate is therefore the ratio between the change in population between 1970 and 1980 not accounted for by fertility and migration, divided by 1970 population.[13]

From IPEA we obtain the following 1970 AMC-level control variables, which we use in all our regressions: average years of schooling of individuals aged 25 and above, literacy rates for people 15 years and older, total population, the percentage of people living in urban areas, and life expectancy. We also obtain a set of time invariant AMC-level controls such as latitude, longitude, distance to the state and federal capitals, and an indicator for whether the AMC is located on the coast. For outcomes we got the mean per capita family income at the AMC-level for 1980, 1991 and 2000, which was itself based on the corresponding population censuses. We calculate our main outcome

---

[13]This also includes individuals who emigrated from an AMC but do not report their municipality of origin.

variable, the growth in AMC mean family income per capita, as the log difference in real mean per capita family income between 2000 and 1970. We also obtained several measures of educational attainment in 1980 at the AMC-level, which we also use as outcomes, and which are based on the educational level of individuals 25 years and older. Specifically we use average years of education, and the share of individuals with less than 4, between 4 and 8, and with 8 or more years of education for each AMC.

Other IPEA data include the value of capital stocks held by businesses in each AMC in 1970 and 1980 in the agricultural, commercial, manufacturing and service sectors, all based on the respective economic censuses.[14] Up until 1980, Brazil's statistical agency carried out periodic economic censuses covering all firms in each of these sectors. As explained in detail by the academics in charge of performing these calculations at IPEA (Reis et al., 2005), when calculating the value of capital stocks for agricultural establishments they include farmland, buildings, long-term crops,[15] vehicles, machinery, agricultural instruments, and livestock. They deduct the value of residential buildings within farms, and only consider livestock used for traction or reproduction. Firms in the agricultural sector include all establishments dedicated to farming, cattle, poultry or rabbits, beekeping, raising silk worms, horticulture, floriculture, forestry, and extraction of vegetable products. When calculating the value of capital stocks for manufacturing, commercial and service industry establishments, they take into account the value of firms' capital employed in buildings, land, machinery and equipment as reported in the corresponding economic censuses. The firms covered by the commercial census are all the establishments dedicated to the purchase, sale, exchange or distribution of merchandise through retail.[16] Activities considered in the manufacturing census include the processing and packaging of food products, metallurgical activities, production of pharmaceutical products, clothes items, etc. Finally, firms in the service sector include all establishment whose activity in-

---

[14]A detailed account on how the value of capital stocks at the AMC-level was backed out from the corresponding economic censuses by IPEA can be found in Reis et al. (2005).

[15]Long-term crops are those that do not need to be replanted after each harvest, such as coffee, oranges, bananas, etc.

[16]For example, the sales activities of a firm that produces machinery is accounted for in the commercial census only if the firm sells its products through its own retail establishments, but not if it does so through a wholesaler. Further explanations can be found in the reports by IBGE on the commercial censuses. For example, at http://biblioteca.ibge. gov.br/visualizacao/periodicos/63/cc_1980_v4_n15_ba.pdf.

volves providing services to people, such as hotels, repair shops, restaurants, and so on.[17] After calculating the value of each establishment's capital stocks, IPEA aggregates these figures at the municipality level, separately for each sector. In performing this calculation, they consider an establishment as belonging to a municipality if it is located there. As with all of our income figures, capital stocks are expressed in real terms (in 2000 R$). We compute the growth in the value of AMC capital stocks between 1970 and 1980 as the log difference in the real value of capital stocks in all sectors of the economy.

We summarize the main variables for our analysis in Table 1. In 1970 Brazil was an extremely poor country. The average AMC monthly mean per capita family income in 1970 was 56 R$ (in R$ of 2000), which is approximately 38 US dollars as of March 2016. Inequality rates were high with an average Gini coefficient of 0.47 and standard deviation 0.07. Inequality also displayed a considerable degree of spatial variation across AMCs, as shown in Figure 2. During the 1970-2000 period income per capita more than doubled on average across AMCs. Most of these gains occurred in the first decade and were accompanied by large increases in physical capital stocks across sectors.

4   Estimation Approach

In order to estimate the effect of initial overall income inequality on subsequent economic growth, we run the following OLS regression:

$$ln(\overline{y}_{a,s,t}) - ln(\overline{y}_{a,s,1970}) = \beta_0 + \beta_1 ln(\overline{y}_{a,s,1970}) + \beta_2 Gini_{a,s,1970} + X_{a,s,1970}\delta + \gamma_s + U_{a,s,t} \quad (1)$$

where $\overline{y}_{a,s,t}$ is the mean per capita family income in AMC a in state s and year t. We estimate this regression with growth in AMC mean per capita family income in 1970-2000 as our outcome variable (i.e., when t is 2000). In order to assess the effect of initial inequality on future levels of income per capita directly, we also use the same specification but replace income growth with

---

[17] Further details can be found in IBGE's reports on the results of the service industry census. See http://biblioteca. ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/censodosservicos/1980_v05_n03_AC.pdf.

11

the (natural) logarithm of average per capita family income in 1980, 1991 and 2000, $ln(\overline{y}_{a,s,t})$, as the outcome variable. $Gini_{a,s,1970}$ is the Gini coefficient in AMC a in state s in 1970, $X_{a,s,1970}$ is a vector of 1970 AMC-level controls, γ_s are state fixed effects and $U_{a,s,t}$ is the influence of unobserved factors on outcomes in year t.[18]

Our coefficient of interest in these regressions is β_2, the effect of initial inequality on the long-run level or growth rate of income per capita. There are many potential confounders at the AMC-level in 1970 that could correlate with both initial income inequality and subsequent economic growth, and the direction of the bias in $\hat{\beta}_2$ is unclear. For instance, AMCs with greater income inequality in 1970 might also be places where a higher percentage of the population has low levels of education, and low education is likely bad for economic growth, biasing $\hat{\beta}_2$ downwards. AMCs with high inequality in 1970 might also be more rural, and growth patterns of rural areas might be different from those of more urbanized AMCs for reasons unrelated to the society's initial income inequality. We address potential omitted variable bias by including standard growth determinants in all our regressions as well as state fixed effects.[19] In particular, $X_{a,s,1970}$ includes a set of AMC characteristics in 1970 (average years of schooling, literacy rate, population, % of urban population, and life expectancy), as well as other time invariant features of each AMC (latitude, longitude, distance from the federal and state capitals, and an indicator for whether the AMC is located on the coast). The key assumption for causal interpretation of $\hat{\beta}_2$ is that unobserved determinants of future income per capita are mean-independent of inequality, conditional on initial income per capita and our controls. While this assumption is not directly testable, we think it is a reasonable assumption, given our supporting evidence on mechanisms and the fact that our controls absorb a large part of the variation in outcomes as shown below.

In order to distinguish between effects of inequality originating from either tail of the income distribution, we take advantage of the fact that the Gini coefficient can be approximated with a

---

[18]Controlling for AMC fixed effects would require us to find valid instruments to deal with the presence of the lagged outcome. We prefer the simplicity and transparency of standard regression control. Moreover, as pointed out by Easterly (2007) for example, it is not clear whether using relatively high-frequency data is appropriate for the question we wish to study since income inequality tends to be fairly stable over time and its impact on subsequent growth is believed to unfold over decades, if not generations. Our estimates are consistent with this notion: although the effect of initial income inequality on growth is strongest in the first decade, it takes around 20 years for the full effect to materialize.

[19]Excluding *Distrito Federal* which is also a municipality in itself, Brazil has 26 states in total.

formula based on the shares of income held by each of the quintiles. That is:

$$Gini_{a,s,1970} \approx 0.8 \times [Q5_{a,s,1970} + 0.5Q4_{a,s,1970} - 0.5Q2_{a,s,1970} - Q1_{a,s,1970}] \tag{2}$$

where $Qn_{a,s,1970}$ is the 1970 share of total income of AMC a in state s held by quintile n. As can be seen in the first column of Table 2, controlling for state fixed effects and our vector of 1970 AMC covariates, the Gini coefficient and its approximation based on quintile shares in 1970 vary almost one-to-one, with an R^2 of almost one. In light of this, decomposing differences in the 1970 AMC Gini coefficients into differences in quintile income shares as in (2) allows us to differentially focus on the growth effects of inequality in the left and right tails of the income distribution. Throughout our Gini decomposition exercise, the omitted quintile is the middle one. Thus, a decrease in the income held by the first quintile implies an increase in the income held by the middle one, and a higher overall income inequality, as illustrated in Panel A of Figure 1. Throughout the paper we refer to this as inequality in the left or bottom tail. The exact same increase in overall inequality occurs when a higher percentage of overall income is held by the top quintile at the expense of the middle one. This is what we call higher inequality in the right or upper tail, as shown in Panel B of Figure 1. With this intuition in mind, we distinguish between the growth effects of inequality in the left and right tails by running the following regressions:

$$ln(\overline{y}_{a,s,t}) - ln(\overline{y}_{a,s,1970}) = \beta_0 + \beta_1 ln(\overline{y}_{a,s,1970}) + \alpha_1 Q5_{a,s,1970} + \alpha_2 Q4_{a,s,1970} + $$
$$+ \alpha_3 Q2_{a,s,1970} + \alpha_4 Q1_{a,s,1970} + X_{a,s,1970}\delta + \gamma_s + U_{a,s,t} \tag{3}$$

which is the specification in (1), but replacing the Gini coefficient with four of the quintile income shares and omitting the middle quintile share. In this regression, our coefficients of interest are α_1 (the coefficient for inequality in the right tail), and α_4 (the coefficient for inequality in the left tail, when multiplied by minus 1). When exploring the correlation between inequality in the left and right tails with subsequent growth in physical capital, we run (3) with the real value of the aggregate capital stock held by firms instead of per capita income. We do this separately for each sector of the

economy (agriculture, manufacturing, commerce and services), and also for the total capital stock across sectors. In all these regressions we control for the log of the 1970 value of the capital stocks held by firms in all sectors (as well as the log of the respective sector-specific capital stocks in 1970).

When analyzing growth in human capital, we run the above regression for a set of outcomes capturing the 1980 levels educational attainment in an AMC, such as average years of education of individuals above 25 years old, the percent of such individuals with less than 4 years of education (i.e, less than a primary school degree), between 4 and 8 years (i.e., more than primary but less than middle school), and 8 or more years of education (i.e., at least a middle school diploma). In addition to the baseline controls included in $X_{a,s,1970}$, we also control for the 1970 proportion of individuals 25 and older with less than 4, between 4 and 8, and 8 or more years of education.

5 Results

5.1 Overall Inequality and Income per Capita Growth

Column 1 of Table 3 shows that there is positive correlation between the Gini coefficient in 1970 and growth in the period spanning 1970-2000. AMCs with a Gini that was one standard deviation (0.07 Gini points) higher in 1970 grew about 3% more between 1970-2000. The results are very similar when using the Gini approximation based on quintile shares, as can be seen in column 2, lending credibility to the regressions based on equation (3) below. Moving to the regressions with the log of income per capita in 1980, 1991 and 2000 as outcomes, it is clear from columns 3 to 5 that the results are much stronger in the first decade. Income per capita increased by about 2% by 1980 and by about 3% by 1991 in places where the Gini coefficient was 7 percentage points higher, with only negligible additional growth by the year 2000. Taking these results together, we conclude that AMCs with higher inequality in 1970 end up with higher average income in 2000, but do not experience permanently higher growth.

In line with the results from cross-country growth regressions, the coefficient for the income lag is negative and statistically significant in the first two columns, meaning that AMCs that start

out with a higher income level grow at a slower rate. Even though our study explores within-country (across sub-national unit) variation, Brazilian AMCs also experience income convergence as predicted by growth theories, which speaks to the external validity of our study and suggests that there are at least some common mechanisms linking inequality and growth both within and across countries. We also note that our regressions account for most of the variation in subsequent income per capita levels, (R^2 of 0.877 in column 5 of Table 3), leaving little room for unobserved confounders to dramatically alter our estimate of interest.

5.2 Quintile Income Shares and Income per Capita Growth

Having established a positive correlation between inequality as measured by the Gini coefficient in 1970 and subsequent economic growth, we now explore whether this effect is different when inequality originates from the lower as opposed to the upper tail of the income distribution. As explained in Section 4, the third quintile is omitted in our regressions with quintile shares. Thus, a *decrease* in the income held by the first quintile is matched by an equivalent increase in the income share of the middle quintile, implying *higher* inequality in the left tail. Therefore, multiplying the coefficient associated with the share of 1970 AMC income held by Q1 by -1 gives us the the effect of an increase in left-tail inequality on growth of AMC income per capita. On the other hand, an *increase* in the share of income held by the top quintile implies a decrease in the share of income held by the middle quintile, and an *increase* in inequality in the right tail. Therefore, the coefficient on the share of 1970 AMC income held by Q5 directly gives the partial effect of an increase in inequality in the right tail on growth of AMC income per capita.

As shown in the first column of Table 4, AMCs with higher inequality in the right tail (i.e., a higher share of income held by Q5 at the expense of Q3) did not grow more in 1970-2000. On the other hand, the negative coefficient for the share of income held by Q1 means that AMCs with higher inequality in the left tail of the distribution did experience higher growth. In particular, AMCs with a 1970 Q3 income share higher by one standard deviation (3 percentage points) at the expense of Q1 grew about 3% more over the period 1970-2000. Income per capita increased by about 4% by 1980

with little additional growth by 1991 and a slight and statistically insignificant drop by 2000. As with overall inequality, higher left-tail inequality does not lead to a permanent increase in income per capita growth. The last row of Table 4 shows that the differential impact between left- and right-tail inequality is not only economically but also statistically significant in most specifications. We therefore conclude that the overall effect of inequality picked up by the Gini coefficient is entirely driven by the lower tail of the initial income distribution: compared to more equal places, AMCs with a higher share of income going to the middle quintile at the expense of the bottom quintile grow more rapidly, while places with a higher share of income going to the top quintile at the expense of the middle quintile get no growth boost at all.

5.3   Quintile Income Shares and Subsequent Income Distribution

In this subsection, we analyze whether the higher growth in per capita income experienced by AMCs with greater initial inequality in the left tail had any distributional impacts. Since this higher growth already materializes by 1980, we focus on this period. As displayed in Table 5 and Figure 3, more inequality in the left tail in 1970 is correlated with a positive shift in the top half of the AMC income distribution. More specifically, in AMCs in which the share of income held by Q3 (Q1) was 3 percentage points higher (lower) in 1970, the per capita income percentiles in the top half of the distribution were between 4% and 6% higher in 1980.

Consistent with these results, Table 6 shows that higher initial inequality in the left tail is associated with significantly lower poverty rates, but only for broad definitions of poverty. Under our two broadest definitions, for which the average poverty rates were 60% and 45%, AMCs in which the share of income held by Q3 in 1970 was 3 percentage points higher (and the share of Q1 was lower) had a poverty rate of about 1 percentage point lower in 1980. Higher initial left-tail inequality does not correlate with lower poverty rates in 1980 for our strict definition of poverty.

6 Evidence on Mechanisms

6.1 Quintile Income Shares and Physical Capital Growth

Given the positive correlation between 1970 inequality in the left tail and subsequent growth in mean per capita family income, we should observe a similar correlation with growth in physical and human capital if credit constraints and setup costs are important. In Table 7, we regress real growth in the value of the capital stocks held by firms from different sectors in the period spanning 1970-1980 against the quintile income shares, the 1970 value of both total and sector-specific capital stocks in 1970 and our other controls. Consistent with the zero growth effect of right-tail inequality discussed above, we find small and statistically insignificant effects of inequality in the right tail on growth in firms' capital stocks for three out of four sectors as well as overall. On the other hand, we find a positive and sizable correlation between inequality in the left tail in 1970 and growth in the value of capital stocks from 1970 to 1980 for all four sectors as well as overall. Total capital stocks grew about 10% more in real terms in AMCs in which the 1970 income share of Q3 was one standard deviation higher at the expense of the bottom quintile. The effect of left-tail inequality on physical capital accumulation arises across sectors, ranging from about 9% in agriculture, to about 16% in the commercial sector, about 27% in manufacturing and about 15% in services.

6.2 Quintile Income Shares and Human Capital Growth

Turning to investments in education, Table 8 shows a similar pattern. The first column documents that there is no correlation between initial inequality in the right tail and growth in average educational attainment over the 1970-1980 period. On the other hand, average years of schooling experienced higher growth in 1970-1980 in AMCs that started out with higher inequality in the left tail. Though significant statistically, this effect is relatively small: AMCs in which the income held by the middle quintile in 1970 was 3 percentage points higher (at the expense of the bottom quintile) saw an increase of 0.03 years in average educational attainment of individuals above 25 years of age. Turning to the regressions in columns 2 to 4, it is clear that this increase in educational attainment

17

was driven by a smaller proportion of the population with less than 4 years of education (i.e., less than a primary school degree), and a higher proportion with educational attainment of between 4 and 8 years. Higher inequality at the top also increased the proportion of the population with more than 8 years of schooling, but the impact is negligibly small.

6.3 Quintile Income Shares and Migration

A plausible concern is that our results are driven by differential migration patterns. As shown in column 1 of Table 9, AMCs with a lower Q1 income share and higher share of income held by the third quintile exhibit higher population growth between 1970 and 1980, but the impact is very small. In particular, the population in AMCs with a 1970 Q3 income share higher by one standard deviation (3 percentage points) at the expense of Q1 grew about 4.5% more over the period 1970-1980. One concern is that AMCs with high initial left-tail inequality could be attracting workers with higher education and higher potential earnings, leading to a selection-driven increase in average income. However, as shown in column 2 of Table 9, this increase in population was not driven by immigration, since AMCs with higher inequality in the left tail in 1970 did not experience higher immigration between 1970 and 1980, the decade of highest growth. Alternatively, places with high initial inequality in the left tail might experience higher out-migration of the poor and thus higher income per capita among remaining residents in future periods. As shown in column 3 of Table 9, we actually find the opposite: AMCs with higher inequality in the left tail in 1970 experienced lower emigration rates.[20] However, this effect is negligible in practice. A 3 percentage point increase in the share of total income held by the middle quintile (at the expense of the bottom quintile) is associated with an emigration rate 0.70 percentage points lower over the 1970-1980 decade relative to an average emigration rate of about 19%. It turns out that the higher growth in population between 1970 and 1980 experienced by AMCs that started out with higher inequality in the left tail is mostly driven by a lower mortality rate, as shown in column 5 of Table 9.[21]

---

[20]These results should be taken with care, since the municipality of origin is missing for approximately 19% of all immigrants.

[21]What we refer to as mortality rate is actually a residual category, namely the ratio between the change in population between 1970 and 1980 not accounted for by fertility and migration and the 1970 population. This includes not only

7 Robustness Checks

7.1 Controlling for 1970 Sectoral Labor Force Shares

While our main specification controls for the share of an AMC's 1970 population living in rural areas, as a robustness check we also account for differences in the initial structure of the economy in a more flexible manner. As detailed in Section 3, we control for the share of occupied individuals working in each of the 16 economic sectors defined by the 1970 census. As shown in Tables A.1 and A.2, the association between inequality in 1970 and subsequent economic growth is robust to the inclusion of these controls. Our evidence on channels featured in Tables A.3 and A.4 is also consistent with our results on inequality, although slightly weaker when it comes to physical capital accumulation.

7.2 Imputing Top-Coded Incomes

Unlike subsequent censuses, incomes in the 1970 census are top-coded, a practice which affects 0.04% of employed individuals. In order to check whether our results are driven by differential measurement error at the bottom versus at the top of the initial income distribution, we impute top-coded incomes and construct new quintile shares. Following the methodology used by Katz and Murphy (1992), Autor et al. (2008), and Autor and Dorn (2013), among others, we multiply top-coded incomes by a factor of 2.15, so that individual incomes in the top 20% follow a Pareto distribution. As can be seen in Table A.5, our main results are robust to these imputations.

7.3 Alternative Definition of the 1970 Census Universe

As explained in Section 3, the 1970 AMC income statistics used in our main specification exclude individuals living in collective dwellings and individuals who live in a private dwelling but are unrelated to the family head (i.e., tenants and domestic servants), which in total account for 4.10%

---

people who passed away in 1970-1980, but also individuals who emigrated from the AMC but did not report their municipality of origin.

of individuals in the 1970 census. While the correlation between initial inequality in the left tail and subsequent growth in income per capita is robust to the inclusion of these individuals, as shown in Table A.6, inequality at the top is positively correlated with growth in income per capita in some specifications. However, the coefficients for inequality in the top are smaller and not robust across specifications.

7.4   Adjusting for Selection on Unobservables

As discussed in Section 4, there could be many confounders at the AMC-level in 1970 correlating with both initial income inequality and subsequent economic growth. Although we address potential omitted variable bias by including standard growth determinants in all our regressions as well as state fixed effects, we cannot fully rule out the existence of unobservable determinants of AMC growth which correlate with initial income inequality even conditional on these controls.

In this subsection, we follow the approach of Oster (Forthcoming), itself an extension of the methodology developed by Altonji et al. (2005), to evaluate the robustness of our estimates to potential omitted variable bias. Under the two assumptions that observable and unobservable variables are equally related to the regressor of interest and that the bias from unobservables is not so large that it biases the direction of the covariance between the observables and the regressor of interest, Oster (Forthcoming) develops an estimator that accounts for selection on unobservables. We also assume, as proposed in that paper, that the hypothetical maximum R^2 from a regression of the dependent variable against all observable and unobservable controls is the minimum value between 1 and 1.3 times the R^2 of the regression with observable controls. Since the quintile income shares only capture inequality in the left and right tails if they are conditioned on the other quintile shares and initial income, we include all of these in the "uncontrolled" regression.

Our estimates are almost unchanged if we adjust them to account for potential selection on unobservables. The bias-adjusted estimate for the Gini coefficient in the regression using growth in 1970-2000 as the outcome variable is equal to 0.402, down from 0.447 in the specification that controls for all our observables (column 1 in Table 3). Intuitively, the change is so small because the

increase in R^2 is massive (from 0.013 in the uncontrolled regression to 0.673 controlling for observables) compared to the change in coefficient estimates (from an uncontrolled 0.541 to a regression-controlled 0.447). If we apply the same procedure to the first quintile income share (again with growth in 1970-2000 as the outcome variable), the adjusted estimate is -0.902, which is practically unchanged from the controlled estimate of -1.065 in column 1 of Table 4. In the case of Q5, which was small and statistically insignificant in our initial regression, controlling for potential omitted variable bias results in an impact estimate of 0.169, compared to 0.243 in column 1 of Table 4.

8 Conclusion

This study investigates whether inequality originating in the lower as opposed to the upper tail of the income distribution has different effects on subsequent income per capita growth. Using subnational variation in Brazil, we find that holding average income per capita and an extensive set of controls constant, AMCs with higher inequality in the left tail of the income distribution in 1970 exhibited higher growth in income per capita over the subsequent three decades. At the same time there is no correlation between initial inequality in the right tail of the AMC income distribution and growth. We show that our estimates are remarkably robust when we account for selection on unobservables. Moreover, our results are barely affected if we flexibly control for 1970 structural differences across sectors, impute incomes that were top-coded in the 1970 census, or use alternative definitions of the population underlying our inequality measures. Consistent with the existence of credit constraints and setup costs for investing in physical and human capital, we show that AMCs that started out with higher inequality in the left tail also accumulated physical and human capital at a faster pace while right-tail inequality has no such effects.

Whether higher left-tail inequality would lead to higher growth in other contexts depends on the level of income and setup costs. Consider once more a stylized economy in which the population is divided into three groups of equal size (the poor, the middle class, and the rich). If the incomes of the poor and the middle class are initially too low to overcome the setup costs for investing in either human or physical capital, an increase in inequality at the bottom while keeping overall

income constant (i.e., a transfer of income from the poor to the middle class) might allow the middle class to overcome the setup costs and invest in human and physical capital. But consider instead an economy where credit constraints only bind for the poorest group. Higher inequality in the left tail would have no impact on growth in this situation. In an even richer economy in which all groups can profitably invest in human and physical capital, higher inequality in the lower tail could even be bad for growth if it results in the poor becoming credit constrained. Assessing the external validity of the credit constraint cum setup cost mechanism thus requires direct evidence on human and physical capital accumulation in line with whatever relationship between left- or right-tail inequality and growth is found in a given setting.

References

Aghion, Philippe and Patrick Bolton, "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 1997, *64* (2), 151–172.

Alesina, Alberto and Dani Rodrik, "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 1994, pp. 465–490.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 2013, *103* (5), 1553–1597.

__ , **Lawrence F. Katz, and Melissa S. Kearney**, "Trends in US Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Banerjee, Abhijit V. and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, *101* (2), 274–298.

Barro, Robert J., "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth*, 2000, *5* (1), 5–32.

Benabou, Roland, "Inequality and Growth," in "NBER Macroeconomics Annual 1996, Volume 11," MIT Press, 1996, pp. 11–92.

Benjamin, Dwayne, Loren Brandt, and John Giles, "Did Higher Inequality Impede Growth in Rural China?," *Economic Journal*, 2011, *121* (557), 1281–1309.

Bourguignon, Francois, "Pareto Superiority of Unegalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function," *Econometrica*, 1981, pp. 1469–1475.

Campante, Filipe R. and Francisco H.G. Ferreira, "Inefficient Lobbying, Populism and Oligarchy," *Journal of Public Economics*, 2007, *91* (5), 993–1021.

Deininger, Klaus and Lyn Squire, "New Ways of Looking at Old Issues: Inequality and Growth," *Journal of Development Economics*, 1998, *57* (2), 259–287.

Easterly, William, "Inequality Does Cause Underdevelopment: Insights from a New Instrument," *Journal of Development Economics*, 2007, *84* (2), 755–776.

Esteban, Joan and Debraj Ray, "Wealth Constraints, Lobbying and the Efficiency of Public Allocation," *European Economic Review*, 2000, *44* (4), 694–705.

Ferreira, Francisco H. G., Christoph Lakner, Maria Ana Lugo, and Berk Özler, "Inequality of Opportunity and Economic Growth: A Cross-Country Analysis," 2014.

Forbes, Kristin J., "A Reassessment of the Relationship between Inequality and Growth," *American Economic Review*, 2000, pp. 869–887.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 1993, *60* (1), 35–52.

_ **and Omer Moav**, "From Physical to Human Capital Accumulation: Inequality and the Process of Development," *Review of Economic Studies*, 2004, *71* (4), 1001–1026.

Hagopian, Frances, *Traditional Politics and Regime Change in Brazil*, Cambridge University Press, 1996.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Li, Hongyi, Lyn Squire, and Heng-Fu Zou, "Explaining International and Intertemporal Variations in Income Inequality," *Economic Journal*, 1998, *108* (446), 26–43.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business Economics and Statistics*, Forthcoming.

Panizza, Ugo, "Income Inequality and Economic Growth: Evidence from American Data," *Journal of Economic Growth*, 2002, *7* (1), 25–41.

Partridge, Mark D., "Is Inequality Harmful for Growth? Comment," *American Economic Review*, 1997, *87* (5), 1019–1032.

Perotti, Roberto, "Growth, Income Distribution, and Democracy: What the Data say," *Journal of Economic Growth*, 1996, *1* (2), 149–187.

Persson, Torsten and Guido Tabellini, "Is Inequality Harmful for Growth?," *American Economic Review*, 1994, pp. 600–621.

Ravallion, Martin, "Why Don't We See Poverty Convergence?," *American Economic Review*, 2012, *102* (1), 504–523.

__ **, Shaohua Chen, and Prem Sangraula**, "Dollar a Day Revisited," *World Bank Economic Review*, 2009, p. lhp007.

Reis, Eustáquio, Kepler Magalhães, Márcia Pimentel, and Mérida Medina, "Estoque de Capital Privado nos Município Brasileiros, 1970-85," *Rio de Janeiro, IPEA*, 2005.

Voitchovsky, Sarah, "Does the Profile of Income Inequality Matter for Economic Growth?," *Journal of Economic Growth*, 2005, *10* (3), 273–296.

Table 1: Descriptive Statistics

	Mean	Std. Dev	Min	Max
Dependent Variables				
Ln (real mean per capita family income)				
1970-2000 growth (log difference)	1.13	0.35	-0.75	3.67
1980 mean (in 2000 R$)	4.77	0.59	2.32	6.41
1991 mean (in 2000 R$)	4.64	0.59	3.17	6.38
2000 mean (in 2000 R$)	5.01	0.57	3.62	6.86
Growth in aggregate capital stocks in 1970-1980				
Agriculture	1.32	0.70	-2.99	18.72
Commerce	1.28	1.17	-13.02	10.41
Manufacturing	1.66	2.47	-12.55	20.36
Services	2.00	2.11	-12.62	16.59
Total	1.39	0.61	-1.40	7.09
1980 educational attainment (people 25 years and older)				
Average years of schooling	2.07	1.06	0.10	7.20
Proportion with less than 4 years of schooling	0.74	0.16	0.15	0.99
Proportion with 4 or more and less than 8 years of schooling	0.20	0.12	0.01	0.75
Proportion with 8 or more years of schooling	0.06	0.05	0.00	0.48
Explanatory Variables - all measured in 1970				
Gini coefficient	0.47	0.07	0.25	0.97
Gini approximation based on quintile income shares	0.42	0.07	0.15	0.80
Share of AMC income held by Q1	0.05	0.02	0.00	0.14
Share of AMC income held by Q2	0.09	0.02	0.00	0.30
Share of AMC income held by Q3	0.14	0.03	0.00	0.27
Share of AMC income held by Q4	0.20	0.03	0.00	0.45
Share of AMC income held by Q5	0.52	0.07	0.14	1.00
Ln (real mean per capita family income) (2000 R$)	3.89	0.54	0.57	5.70
Average years of schooling (25 years and older)	1.37	0.81	0.00	5.60
Iliteracy rate (15 years and older)	0.44	0.18	0.03	0.92
Population (in 000s)	25.45	132.47	0.83	5924.61
Proportion of residents living in urban areas	0.33	0.21	0.01	1.00
Life expectancy	51.11	4.27	38.40	64.46

Notes: The unit of observation is an Área Mínima Comparável (AMC) over the period 1970-2000. There are 3,659 AMCs.

Figure 1: Gini Coefficient and Inequality at the Top and Bottom of the Income Distribution

Panel A: Higher Overall Inequality Originating from the Left Tail of the Distribution
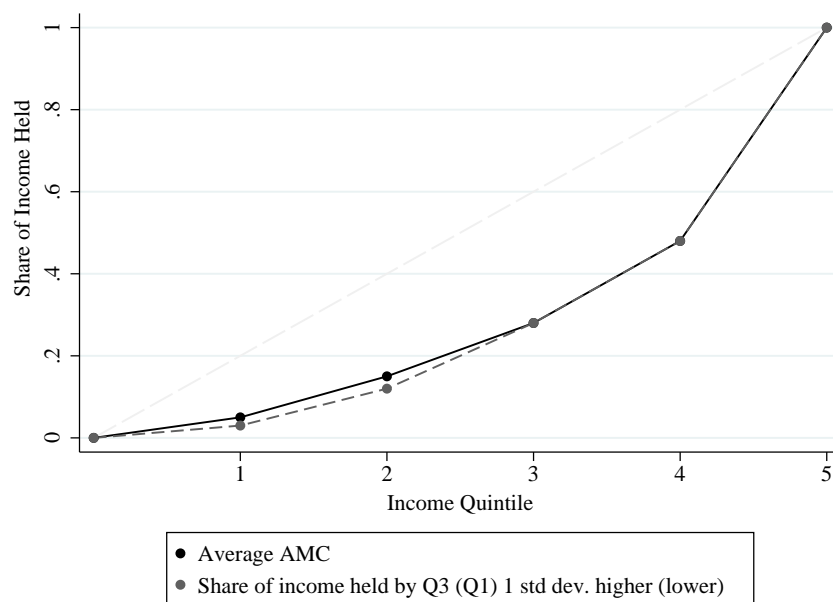
Panel B: Higher Overall Inequality Originating from the Right Tail of the Distribution
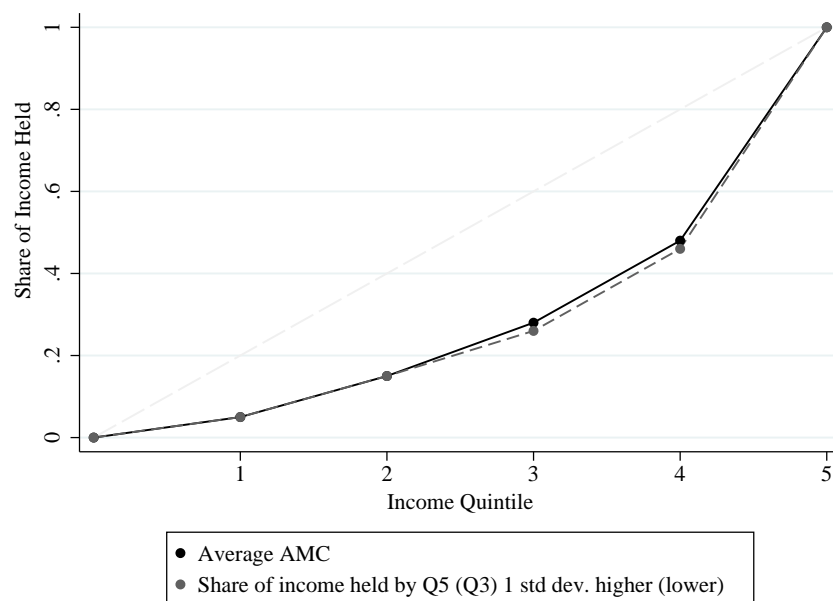
Notes: The solid lines display the Lorenz curve of an AMC with average quintile shares in 1970. *Áreas Mínimas Comparáveis* (AMCs) are roughly equivalent to Brazil's municipalities in 1970. In Panel A, the dashed line shows the Lorenz curve of another AMC where the share of income held by the third quintile is 1 standard deviation (3 percentage points) higher, at the expense of the first quintile, holding the other quintile shares constant. In Panel B, the dashed line shows the Lorenz curve of yet another AMC where the share of income held by the fifth quintile is 1 standard deviation (3 percentage points) higher, at the expense of the middle quintile, again holding the other quintile shares constant.

Figure 2: Gini Coefficient Across Brazilian AMCs in 1970

(0.532, 0.973]
(0.487, 0.532]
(0.448, 0.487]
(0.407, 0.448]
[0.249, 0.407]
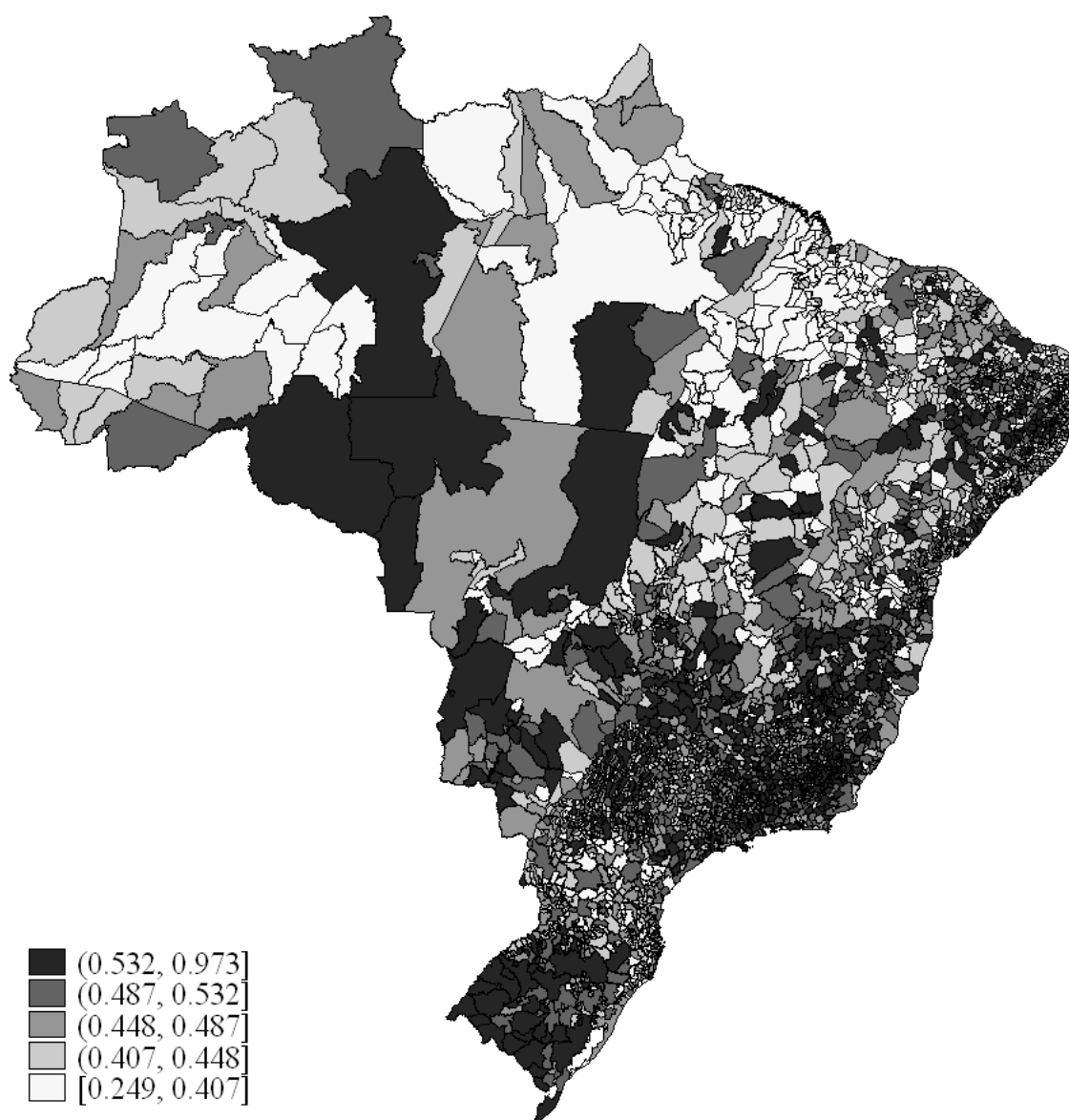
Notes: Each unit is an Área Mínima Comparável (AMC) in 1970. Darker areas indicate greater income inequality as measured by the Gini coefficient in 1970.

Table 2: Income Shares and Income Inequality in 1970

Dependent variable: 1970 Gini coefficient

Gini approximation	1.099***	
	(0.006)	
Share of 1970 AMC income held by Q5		0.925***
		(0.015)
Share of 1970 AMC income held by Q4		0.350***
		(0.023)
Share of 1970 AMC income held by Q2		-0.332***
		(0.026)
Share of 1970 AMC income held by Q1		-0.696***
		(0.029)
Observations	3,659	3,659
R^2	0.976	0.979

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC, and the dependent variable in both columns is the 1970 Gini coefficient. The explanatory variable of interest in column 1 is the 1970 Gini approximation based on quintile shares, calculated using the formula in (2). The explanatory variables of interest in column 2 are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 3: Income Inequality in 1970 and Subsequent Economic Growth

	1970-2000 growth		Ln (Income)		
	Ln difference		1980	1991	2000
Gini coefficient	0.447***		0.313***	0.415***	0.447***
	(0.060)		(0.073)	(0.072)	(0.060)
Gini approximation		0.521***			
		(0.066)			
Ln (1970 income)	-0.771***	-0.770***	0.386***	0.296***	0.229***
	(0.017)	(0.017)	(0.023)	(0.021)	(0.017)
Observations	3,659	3,659	3,659	3,659	3,659
R^2	0.673	0.673	0.857	0.849	0.877

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in columns 1 and 2 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The 1970 Gini approximation based on quintile shares is calculated using the formula in (2), and Ln (1970 income) is the mean per capita family income in 1970 (in ln). All regressions include state fixed effects and control for average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 4: Income Shares in 1970 and Subsequent Economic Growth

	1970-2000 growth	Ln (Income)		
	Ln difference	1980	1991	2000
Share of 1970 AMC income held by Q5	0.243	0.183	0.092	0.243
	(0.170)	(0.226)	(0.198)	(0.170)
Share of 1970 AMC income held by Q4	0.216	0.126	-0.026	0.216
	(0.229)	(0.328)	(0.255)	(0.229)
Share of 1970 AMC income held by Q2	-0.349	0.359	-0.125	-0.349
	(0.277)	(0.336)	(0.320)	(0.277)
Share of 1970 AMC income held by Q1	-1.065***	-1.440***	-1.591***	-1.065***
	(0.380)	(0.469)	(0.453)	(0.380)
Observations	3,659	3,659	3,659	3,659
R^2	0.674	0.858	0.849	0.877
P-value (Q4+Q2=0)	0.746	0.362	0.748	0.746
P-value (Q5+Q1=0)	0.110	0.055	0.014	0.110

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in column 1 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 5: Income Shares in 1970 and 1980 Income Percentiles

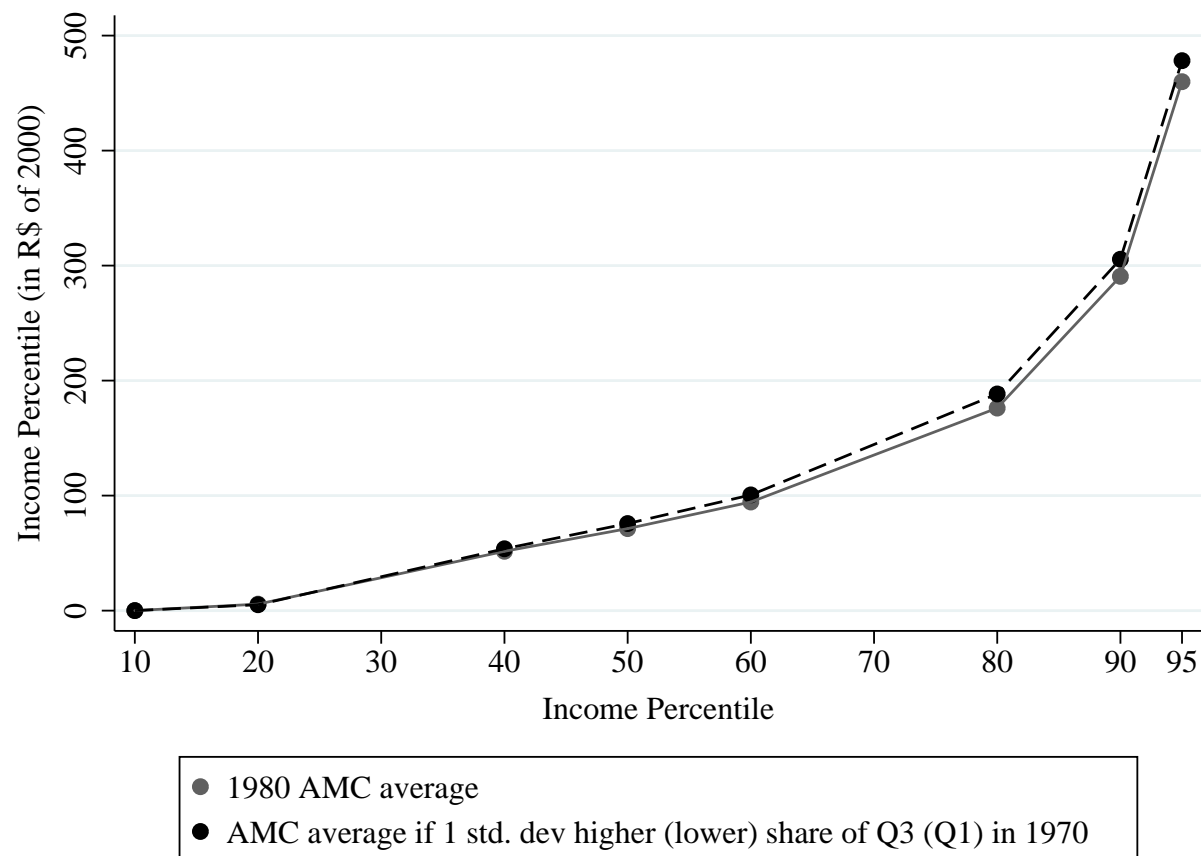|	Ln (Income Percentile in 1980)							
	10th	20th	40th	50th	60th	80th	90th	95th
Share of 1970 AMC income held by Q5	-0.038	0.262	-0.421	-0.576	-0.790**	-0.655*	-0.123	0.382*
	(0.028)	(0.688)	(0.455)	(0.423)	(0.382)	(0.341)	(0.196)	(0.216)
Share of 1970 AMC income held by Q4	-0.043	0.713	0.278	0.097	-0.148	-0.407	0.050	0.339
	(0.065)	(0.921)	(0.662)	(0.626)	(0.570)	(0.593)	(0.274)	(0.305)
Share of 1970 AMC income held by Q2	-0.067	3.425***	1.160	0.937	0.342	0.351	0.101	0.134
	(0.051)	(1.165)	(0.827)	(0.727)	(0.691)	(0.312)	(0.295)	(0.329)
Share of 1970 AMC income held by Q1	-0.084	2.986**	-1.429	-2.056**	-2.173***	-2.316***	-1.707***	-1.325***
	(0.060)	(1.482)	(0.965)	(0.829)	(0.663)	(0.598)	(0.425)	(0.474)
Observations	3,658	3,658	3,658	3,658	3,658	3,658	3,658	3,658
R^2	0.807	0.415	0.665	0.726	0.791	0.856	0.867	0.854

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variables are the different AMC income percentiles in 1980 (in ln), based on per capita family incomes. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Figure 3: Inequality in the Left in 1970 and 1980 Income Percentiles

Notes: The solid line plots the 1980 income percentiles of the average AMC in terms of per capita family income. The dashed line plots the 1980 income percentiles of an AMC with a 1 standard deviation higher (lower) share of income held by the third (first) quintile in 1970, which was calculated using the coefficients in Table 5.

Table 6: Income Shares in 1970 and 1980 Poverty Rates

	% of people under poverty line in 1980		
	1/2 the Sep-91 min. wage (84.73 R\$ a month)	US\$ 2 a day (50.67 R\$ a month)	US\$ 1.25 a day (26.43 R\$ a month)
Share of 1970 AMC income held by Q5	0.158***	0.084	0.014
	(0.057)	(0.061)	(0.062)
Share of 1970 AMC income held by Q4	0.004	-0.037	-0.070
	(0.073)	(0.082)	(0.088)
Share of 1970 AMC income held by Q2	-0.188**	-0.205**	-0.140
	(0.090)	(0.104)	(0.102)
Share of 1970 AMC income held by Q1	0.417***	0.326**	0.106
	(0.127)	(0.136)	(0.132)
Observations	3,658	3,658	3,658
R^2	0.871	0.826	0.699
Dependent Variable Mean	0.604	0.447	0.319

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable is the share of people in the AMC in 1980 below the poverty line in terms of their per capita family income. The poverty line used in column 1, obtained from IPEA, is half the Brazilian minimum wage in September 1991, whereas the poverty lines in columns 2 and 3 (US\$ 2 and US\$ 1.25 a day at 2005 PPP) were taken from Ravallion (2012). The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 7: Income Shares in 1970 and Real Growth in Value of Firms' Capital Stocks between 1970 and 1980

	Agriculture	Commercial	Manufacturing	Services	Total
Share of 1970 AMC income held by Q5	-0.052	-1.168	-5.228**	-0.726	-0.857
	(0.596)	(0.868)	(2.052)	(1.231)	(0.538)
Share of 1970 AMC income held by Q4	0.060	-0.647	-5.063*	-0.183	0.121
	(0.751)	(1.231)	(2.599)	(1.735)	(0.730)
Share of 1970 AMC income held by Q2	0.262	0.159	-2.920	-0.504	-0.313
	(1.014)	(1.448)	(3.860)	(2.029)	(1.012)
Share of 1970 AMC income held by Q1	-2.962**	-5.480***	-8.951**	-4.900**	-3.576***
	(1.336)	(1.916)	(3.909)	(2.438)	(1.244)
Observations	3,659	3,659	3,659	3,659	3,659
R^2	0.206	0.240	0.362	0.637	0.109
P-value (Q5+Q1=0)	0.093	0.011	0.010	0.099	0.007

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variables are the 1970-1980 growth rate in the value of the AMC's private sector capital stocks for each productive sector, calculated by IPEA from the 1970 and 1980 economic censuses. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 capital stocks in all sectors, mean per capita family income (in ln), average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 8: Income Shares in 1970 and Educational Attainment in 1980

	Average years of education	% of people by years of education		
		< 4 years	≥ 4 and < 8 years	≥ 8 years
Share of 1970 AMC income held by Q5	0.323	-0.002	-0.024	0.026*
	(0.225)	(0.037)	(0.035)	(0.014)
Share of 1970 AMC income held by Q4	0.390	-0.031	0.001	0.031*
	(0.311)	(0.052)	(0.049)	(0.018)
Share of 1970 AMC income held by Q2	0.388	-0.055	0.062	-0.006
	(0.373)	(0.064)	(0.059)	(0.023)
Share of 1970 AMC income held by Q1	-1.091**	0.177**	-0.148**	-0.029
	(0.491)	(0.079)	(0.075)	(0.029)
Observations	3,659	3,659	3,659	3,659
R^2	0.935	0.925	0.897	0.886
Dependent Variable Mean	2.073	0.742	0.196	0.062

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. All dependent variables are calculated in 1980 for individuals 25 years and older. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 capital stocks in all sectors, mean per capita family income (in ln), average schooling attainment, % of people according to educational attainment groups, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

Table 9: Income Shares in 1970, Population Growth and Migration from 1970 to 1980

	Population Growth	Immigration Rate	Emigration Rate	Fertility Rate	Mortality Rate
Share of 1970 AMC income held by Q5	-0.397	-0.167	0.240***	-0.033	-0.043
	(0.383)	(0.317)	(0.053)	(0.051)	(0.107)
Share of 1970 AMC income held by Q4	-0.004	0.211	0.142*	-0.036	0.037
	(0.427)	(0.356)	(0.073)	(0.064)	(0.135)
Share of 1970 AMC income held by Q2	0.787	0.616	0.032	0.103	-0.101
	(0.661)	(0.466)	(0.100)	(0.105)	(0.223)
Share of 1970 AMC income held by Q1	-1.494**	-0.627	0.232**	-0.150	0.483**
	(0.643)	(0.506)	(0.113)	(0.105)	(0.210)
Observations	3,659	3,659	3,659	3,659	3,659
R^2	0.210	0.194	0.524	0.596	0.276
Dependent Variable Mean	0.137	0.258	0.190	0.242	0.174

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in column 1 is the AMC's population growth rate in 1970-1980, and the dependent variable in column 2 is the immigration rate between 1970 and 1980, i.e., the ratio between the number of people living in the AMC in 1980 who were not living there in 1970 (or who belong to a family in which the head was not living there in 1970 if aged less than 10) and the AMC's population in 1970. The dependent variable in column 3 is the AMC's emigration rate in 1970-1980, calculated as the ratio between the number of people who reported the AMC as their previous residence but were not living there in 1980, and the AMC's population in 1970. The dependent variable in column 4 is the AMC's fertility rate in 1970-1980, computed as the ratio between the number of children less than 10 year old living in the AMC in 1980 whose parents are non-immigrants and the AMC population in 1970. The dependent variable in column 5, which we refer to as mortality rate, is the ratio between the change in population between 1970 and 1980 not accounted for by fertility and migration and the 1970 population; this is a residual category, including not only people who passed away in 1970-1980, but also individuals who emigrated from the AMC but did not report their municipality of origin. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income (in ln), average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.

ONLINE APPENDIX

Appendix Tables

Table A.1: Income Inequality in 1970 and Subsequent Economic Growth (Controlling for Sectoral Labor Shares)

	1970-2000 growth		Ln (Income)		
	Ln difference		1980	1991	2000
Gini coefficient	0.333***		0.291***	0.335***	0.333***
	(0.065)		(0.080)	(0.078)	(0.065)
Gini approximation		0.395***			
		(0.071)			
Ln(1970 income)	-0.789***	-0.788***	0.359***	0.264***	0.211***
	(0.017)	(0.017)	(0.024)	(0.022)	(0.017)
Observations	3,659	3,659	3,659	3,659	3,659
R^2	0.682	0.682	0.860	0.852	0.880

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in columns 1 and 2 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The 1970 Gini approximation based on quintile shares is calculated using the formula in (2), and Ln (1970 income) is the mean per capita family income in 1970 (in ln). All regressions include state fixed effects and control for average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, whether the AMC is located on the coast, and the share of occupied individuals working in the 16 economic sectors defined in the census.

Table A.2: Income Shares in 1970 and Subsequent Economic Growth (Controlling for Sectoral Labor Shares)

	1970-2000 growth	Ln (Income)		
	Ln difference	1980	1991	2000
Share of 1970 AMC income held by Q5	0.053	0.167	-0.026	0.053
	(0.175)	(0.232)	(0.203)	(0.175)
Share of 1970 AMC income held by Q4	0.097	0.118	-0.106	0.097
	(0.228)	(0.332)	(0.254)	(0.228)
Share of 1970 AMC income held by Q2	-0.347	0.345	-0.132	-0.347
	(0.281)	(0.336)	(0.324)	(0.281)
Share of 1970 AMC income held by Q1	-1.244***	-1.378***	-1.622***	-1.244***
	(0.385)	(0.483)	(0.463)	(0.385)
Observations	3,659	3,659	3,659	3,659
R^2	0.683	0.861	0.853	0.881
P-value (Q4+Q2=0)	0.547	0.385	0.616	0.547
P-value (Q5+Q1=0)	0.023	0.073	0.009	0.023

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in column 1 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, whether the AMC is located on the coast, and the share of occupied individuals working in the 16 economic sectors defined in the census.

Table A.3: Income Shares in 1970 and Real Growth in Value of Firms' Capital Stocks between 1970 and 1980 (Controlling for Sectoral Labor Shares)

	Agriculture	Commercial	Manufacturing	Services	Total
Share of 1970 AMC income held by Q5	0.055	-0.952	-3.572*	-0.400	-0.649
	(0.562)	(0.862)	(1.992)	(1.206)	(0.532)
Share of 1970 AMC income held by Q4	0.243	-0.709	-4.404*	-0.343	0.209
	(0.734)	(1.207)	(2.487)	(1.682)	(0.728)
Share of 1970 AMC income held by Q2	0.234	0.521	-2.594	0.205	-0.194
	(0.996)	(1.400)	(3.809)	(1.966)	(1.007)
Share of 1970 AMC income held by Q1	-3.026**	-4.433**	-5.183	-2.742	-3.188**
	(1.295)	(1.917)	(3.913)	(2.447)	(1.256)
Observations	3,659	3,659	3,659	3,659	3,659
R^2	0.230	0.271	0.387	0.651	0.119
P-value (Q5+Q1=0)	0.081	0.039	0.109	0.355	0.020

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variables are the 1970-1980 growth rate in the value of the AMC's private sector capital stocks for each productive sector, calculated by IPEA from the 1970 and 1980 economic censuses. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 capital stocks in all sectors, mean per capita family income (in ln), average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, whether the AMC is located on the coast, and the share of occupied individuals working in the 16 economic sectors defined in the census.

Table A.4: Income Shares in 1970 and Educational Attainment in 1980 (Controlling for Sectoral Labor Shares)

	Average years of education	% of people by years of education		
		< 4 years	≥ 4 and < 8 years	≥ 8 years
Share of 1970 AMC income held by Q5	0.270	0.002	-0.027	0.025*
	(0.230)	(0.037)	(0.035)	(0.013)
Share of 1970 AMC income held by Q4	0.336	-0.026	0.003	0.024
	(0.319)	(0.052)	(0.048)	(0.018)
Share of 1970 AMC income held by Q2	0.533	-0.072	0.061	0.011
	(0.369)	(0.063)	(0.057)	(0.021)
Share of 1970 AMC income held by Q1	-0.947*	0.163**	-0.162**	-0.001
	(0.492)	(0.080)	(0.076)	(0.027)
Observations	3,659	3,659	3,659	3,659
R^2	0.937	0.927	0.899	0.897
Dependent Variable Mean	2.073	0.742	0.196	0.062

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. All dependent variables are calculated in 1980 for individuals 25 years and older. The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 capital stocks in all sectors, mean per capita family income (in ln), average schooling attainment, % of people according to educational attainment groups, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, whether the AMC is located on the coast, and the share of occupied individuals working in the 16 economic sectors defined in the census.

Table A.5: Income Shares in 1970 and Subsequent Economic Growth (Imputing Top-Coded Incomes)

	1970-2000 growth	Ln (Income)		
	Ln difference	1980	1991	2000
Share of 1970 AMC income held by Q5	0.191	0.123	0.033	0.191
	(0.170)	(0.227)	(0.198)	(0.170)
Share of 1970 AMC income held by Q4	0.221	0.127	-0.021	0.221
	(0.229)	(0.330)	(0.256)	(0.229)
Share of 1970 AMC income held by Q2	-0.377	0.324	-0.148	-0.377
	(0.278)	(0.337)	(0.321)	(0.278)
Share of 1970 AMC income held by Q1	-1.117***	-1.494***	-1.651***	-1.117***
	(0.382)	(0.471)	(0.454)	(0.382)
Observations	3,659	3,659	3,659	3,659
R^2	0.677	0.858	0.849	0.877
P-value (Q4+Q2=0)	0.705	0.400	0.722	0.705
P-value (Q5+Q1=0)	0.072	0.037	0.008	0.072

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. The dependent variable in column 1 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast. Top-coded incomes in 1970 are multiplied by a factor of 2.15 so that individual incomes in the top 20% follow a Pareto distribution.

Table A.6: Income Shares in 1970 and Subsequent Economic Growth (Including Collective Households and Non-Family Members)

	1970-2000 growth	Ln (Income)		
	Ln difference	1980	1991	2000
Share of 1970 AMC income held by Q5	0.314**	0.438*	0.267	0.314**
	(0.142)	(0.246)	(0.208)	(0.142)
Share of 1970 AMC income held by Q4	0.419**	0.547	0.326	0.419**
	(0.196)	(0.339)	(0.275)	(0.196)
Share of 1970 AMC income held by Q2	-0.307	0.583	0.014	-0.307
	(0.272)	(0.377)	(0.337)	(0.272)
Share of 1970 AMC income held by Q1	-0.942***	-1.119**	-1.136**	-0.942***
	(0.356)	(0.505)	(0.472)	(0.356)
Observations	3,659	3,659	3,659	3,659
R^2	0.672	0.859	0.849	0.878
P-value (Q5+Q1=0)	0.170	0.338	0.174	0.170

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors are reported in parentheses. The unit of observation is an AMC. For arriving at the AMC per capita family income distribution in 1970 we do not exclude individuals living in collective dwellings and those living in private dwellings which are unrelated to the family head. The dependent variable in column 1 is the 1970-2000 growth in the mean per capita family income, and the dependent variable in columns 3, 4 and 5 is the mean per capita family income in 1980, 1991 and 2000 (in ln). The explanatory variables of interest are the shares of 1970 AMC income held by each of the quintiles, with the third quintile being the omitted category. All regressions include state fixed effects and control for 1970 mean per capita family income, average schooling attainment, literacy rate, population, % of urban population, life expectancy, latitude, longitude, distance from state and federal capital, and whether the AMC is located on the coast.