



HAL
open science

The rich, the poor, and the middle class: banking crises and income distribution

Mehdi El Herradi, Aurélien Leroy

► **To cite this version:**

Mehdi El Herradi, Aurélien Leroy. The rich, the poor, and the middle class: banking crises and income distribution. 2021. halshs-03269859

HAL Id: halshs-03269859

<https://shs.hal.science/halshs-03269859>

Preprint submitted on 24 Jun 2021

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

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

The rich, the poor, and the middle class: banking crises and income distribution

Mehdi El Herradi
Aurélien Leroy

WP 2021 - Nr 36

The rich, the poor, and the middle class: banking crises and income distribution

Mehdi El Herradi* Aurélien Leroy †

June 2021

Abstract

How do banking crises affect rich, middle-class and poor households? This paper quantifies the distributional implications of banking crises for a panel of 140 economies over the 1970–2017 period. We rely on different empirical settings, including an instrumental variable approach, that exploit the geographical diffusion of banking crises across borders. Our results show that banking crises systematically reduce the income share of rich households and positively affect middle-class households. We also find that income inequality increases during periods preceding the triggering of a banking crisis.

JEL Codes: G01, D33, D63

Keywords: Banking crises, income distribution, inequality

*Aix-Marseille Univ, CNRS, AMSE, Marseille. Corresponding author. E-mail: el-mehdi.el-herradi@univ-amu.fr

†University of Bordeaux. E-mail: aurelien.leroy@u-bordeaux.fr.

We thank seminar participants at the 2021 AAP /ANR CaLiBank Workshop for helpful comments. This work was also supported by the French National Research Agency Grants ANR-17-EURE-0020 and the Excellence Initiative of Aix-Marseille University - A*MIDEX.

1 Introduction

The recurrent episodes of financial instability in many advanced and emerging economies since the 1980s have led to questions about the role of finance in our societies. Over the past decades and in a context marked by rising inequality, an important stream of the literature has emerged that reconsiders the impact of finance on income inequality.¹ One dimension that has received attention in the literature is banking crises. In fact, while the relationship between financial development and inequality is still debated, "too much finance" can also be linked to the frequency of banking crises, which in turn affects inequality.² However, the existing literature on the interaction between banking crises and income inequality (i) remains inconclusive, (ii) lacks strong causal assessment and (iii) does not examine the impacts across the total income distribution.

This paper presents new empirical evidence on the distributional consequences of banking crises using annual data from 140 countries spanning the 1970–2017 period. Different segments of the income ladder are considered, consisting of top incomes, the middle class and incomes at the bottom. On the one hand, banking crises primarily induce output losses, but their impacts across the income distribution may be heterogeneous, especially because rich, middle-class and poor households rely on different income sources. On the other hand, there are several databases of banking crises that identify disruptions in the banking system based on exceptional events or policy interventions, such as bank closures and government bailouts. To examine how banking crises affect income distribution, we build on a database offering the largest coverage of banking crisis episodes and introduce original identifications of the causal relationship of interest.

The estimation of the causal effect of a banking crisis on income distribution faces several challenges. First, countries that have experienced a banking crisis differ from those that have not in terms of unobserved economic and institutional characteristics that also have an impact on the dynamics of income distribution. Second, as shown by [Belletini et al. \(2019\)](#), income inequality tends to be higher at the beginning of the crisis than during the years prior to the crisis, which violates the parallel-trends assumption that underlies panel data estimates. Third, although controlling for country fixed effects and income distribution dynamics mitigate these issues, the trigger of a banking crisis could be driven by time-varying unobservable factors, potentially leading to biased estimates. Our paper aims to address these challenges using different settings.

¹See [Bazillier and Hericourt \(2017\)](#) for an excellent review of the relationship between inequality, leverage, and financial crises.

²As discussed by [Loayza et al. \(2018\)](#), financial liberalization is associated with credit expansions and excessive risk-taking, which induces economic fragility and the likelihood of crisis.

The design of our empirical methodology is threefold and broadly draws on [Acemoglu et al. \(2019\)](#). Our first approach infers causal effects from a linear panel model using country fixed effects and controlling for the confounding influence of autoregressive dynamics on income distribution as well as other potential confounding factors. The inclusion of the lags of the income distribution indicators ensures that countries experiencing a banking crisis are not on a differential trend in terms of inequality dynamics. However, the dynamic panel model does not address the possibility that both banking crises and income distribution might be affected by time-varying omitted variables.

Our second approach confronts this issue by using an instrumental variables (IV) approach. The literature emphasizes that banking crises often spread in regional waves (see [van Rijckeghem and Weder \(2001\)](#) and [Dungey and Gajurel \(2015\)](#), among others). We empirically introduce this observation and use regional waves in banking crises as an instrument for banking crises at the country level. Our IV strategy exploits the geographical diffusion of banking crises across countries belonging to the same region with similar economic features. We cautiously ensure that the estimated effect of interest is not driven by unobserved regional factors or common trends in the income distribution at the regional level.

The two previous strategies heavily rely on the linearity assumption and restrict the time pattern of the cumulative effects of a banking crisis on income distribution. That is why our third approach adopts a semiparametric treatment effects framework to examine how the trigger of a banking crisis—the treatment—affects income distribution over time. This approach, which is based on [Angrist and Guido \(2011\)](#) and [Angrist et al. \(2018\)](#), links the selection of a country into a banking crisis to observables, namely, the dynamics of the income distribution and country-level controls. Related to our first approach, this strategy also ensures that countries experiencing a banking crisis are not a differential inequality trend relative to those that have not. Further, it allows us to estimate the dynamic effects of a banking crisis on income distribution.

There is a large body of empirical work aimed at understanding the distributional implications of financial development and financial liberalization. This paper is specifically related to research on the effects of banking crises on income inequality.³ [Morelli \(2018\)](#) analyzes how banking crises in the U.S. affected top incomes over the 1913–1915 period. He finds that systemic banking crises "*reduce income concentration within the top decile of the U.S. pre-tax and transfers income distribution*". From a cross-country perspective, [Roine et al. \(2009\)](#) examine the long-run determinants of top-income shares for 16 countries and document that the outbreak of banking

³Note that there is also a growing literature that examines whether increasing income inequality results in banking crises (see [\(Atkinson and Morelli, 2015; Bellettini et al., 2019; Kirschenmann et al., 2016; Rhee and Kim, 2018\)](#) among others and [van Treeck \(2014\)](#) for a review).

crises is associated with reduced income shares of the rich. [De Haan and Sturm \(2017\)](#) attempt to grasp the entire income distribution for a sample of 121 countries covering 1975–2005 and find that banking crises increase the Gini index, a standard measure of income inequality. However, this evidence is challenged by [Denk and Cournède \(2015\)](#), who find that the effect of banking crises on the Gini index is insignificant for a panel of 31 economies from 1974 to 2011. The same conclusion is reached by [Bazillier and Najman \(2017\)](#) concerning the relationship between banking crises and the labor share, while that related to currency crises is positive. There is also prior work on the consequences of recessions and currency crises on income inequality. For instance, [Cho and Newhouse \(2013\)](#) examines how different types of workers in middle-income countries were affected by the financial crisis of 2007–2008. Their results suggest that female workers and low-skilled workers were not necessarily the most affected during the crisis, while youth experienced significant increases in unemployment and declines in wage employment.

This paper has three main contributions. To the best of our knowledge, this study is the first to provide the most comprehensive country-time coverage on the distributional consequences of banking crises. In this respect, (i) we mobilize [Laeven and Valencia \(2020\)](#)'s database on episodes of banking crises and (ii) exploit the pretax national income shares held by rich, middle-class and poor households, which are obtained from the World Inequality Database (WID). Although the timing of banking crises has been assessed in several works, [Laeven and Valencia \(2020\)](#)'s database seems, according to [Chaudron and de Haan \(2014\)](#), to be more accurate in comparison to other competing databases. In addition, the empirical analysis focuses on different segments of the income distribution to examine the differentiated effects that a banking crisis may produce across the income ladder. Furthermore, we provide new cross-country evidence on the dynamics of income distribution prior to the trigger of banking crises, challenging the previous findings of [Bordo and Meissner \(2012\)](#) and [Atkinson and Morelli \(2011\)](#). Finally, unlike most of the existing evidence on this topic, our analysis aims to support a direct causal relationship between the occurrence of a banking crisis and the segments of the income distribution using an Instrumental Variable (IV) approach.

Our results are easily summarized. We document beforehand that segments of the income distribution feature different dynamics in periods around the banking crisis. In fact, the shares of top incomes are higher at the outbreak of the banking crisis than five years earlier, while those of the middle-class and bottom-income households follow a downward trend. Turning to the empirical estimations, our dynamic panel model shows that the occurrence of a banking crisis is negatively associated with top-income shares and bottom-income shares, but this correlation is positive for the middle class, i.e., households positioned between P21 and P79 of the income ladder. Then, the IV estimates provide insights into the causal effects of banking crises on the

segments of income distribution. As far as rich households are concerned, the trigger of a banking crisis reduces the pre-tax national income share held by the top 1% by 0.77 percentage points. This negative effect also holds for the top 10% and top 20% richest, whose income shares decline by 1.08 and 0.93 percentage points, respectively. Conversely, the middle class is less negatively affected following a banking crisis, leading its national income share to increase by 0.91 percentage points. We also demonstrate that this causal evidence is robust to trends in GDP and the income distribution at the regional level as well as unobserved regional characteristics. Regarding our treatment framework, the resulting estimates successfully control for the influence of the aforementioned patterns in the income distribution. They also establish that the IV results hold over time: banking crises have immediate effects on top incomes and the middle class, and these effects persist in the years following the crisis. In all, these results lead us to confirm that banking crises reduce the income of the rich, while middle-class households are less negatively affected. Nonetheless, the obtained estimates in the three empirical approaches do not seem to indicate a clear-cut relationship between banking crises and bottom-income shares, i.e., the 20% and 10% poorest.

How can these results be interpreted? A plausible explanation concerning the more negative effects on top incomes can be related to the fact that banking crises are typically associated with stock market crashes ([Reinhart and Rogoff \(2013\)](#)). Given that capital incomes constitute the bulk of rich households, it is possible to rationalize the negative effect of banking crises on the top 1%, 10% and 20% following this line of thought. For the middle class, our evidence can support the political economy view of [Chwieroth and Walter \(2017\)](#): banking crises electorally threaten governments in office—especially by middle-class voters—leading them to implement extensive bailouts and other policies aimed at offsetting the consequences of the crisis. Although this applies in particular to democracies, we acknowledge that other stories cannot be ruled out. In this sense, the role of labor market institutions and safety net systems as well as the effect of banking crises on mortgage loans, which are mainly held by middle-class households, should not be ignored.

The rest of this paper is structured as follows. Section 2 describes the data. Section 3 discusses some prima facie evidence on the relationship between banking crises and income distribution. The fourth section introduces the estimation methodologies and the identification strategy. Section 5 presents the results, while the sixth and final section concludes the paper.

2 Data

Our empirical analysis builds on a country-level yearly dataset of 140 countries (Table A1 in the Appendix lists the countries included in the sample) over the 1970–2017 period based on two building blocks, banking crises and the pretax national income shares of different percentiles.

Data on banking crises

Information on the timing of banking crises is obtained from the database of [Laeven and Valencia \(2020\)](#). The authors rely on "events methodology" and define a banking crisis based on two conditions: (i) the banking system should present serious signs of financial distress (significant bank runs, losses in the banking system, bank liquidations) together with (ii) important policy interventions in response to the major disruptions in the banking system.⁴ The crisis variable is one if a country faces a banking crisis and zero otherwise. Several databases on banking crises exist, but they strongly disagree on the start/end of crisis dates, which, consequently, translates into different lengths of banking crises (see [Reinhart and Rogoff \(2009\)](#), for instance). We prefer, however, the [Laeven and Valencia \(2020\)](#)'s database inasmuch as [Chaudron and de Haan \(2014\)](#) show that it is more reliable than competing banking crises databases, in addition to its wide country coverage. Although the simple categorical variable on banking crises prevents us from exploring their different durations and intensities, we adopt an empirical strategy that allows overtime estimation of the effects of a banking crisis on income distribution. Figure A1 in the Appendix provides an overview of the occurrence of banking crises over the studied period and records 459 episodes of banking crises (systemic and nonsystemic).

Data on income distribution

Most previous studies on the crisis-inequality nexus relied on the widely used Gini index, which is a synthetic measure of income inequality. One contribution of this paper is to go beyond the Gini index and examine the distributional effects of banking crises by focusing on different segments of the income ladder. This is done to uncover the differentiated impacts that crises may produce on rich, middle-class and poor households. To achieve this, we use pretax income shares from the World Inequality Database (WID), which combines data from national accounts, household surveys and tax declarations to produce series on the entire distribution of income from the bottom to the top. As explained by [Alveredo et al. \(2016\)](#), the pretax income concept retained in the data is based on the notion of national income (i.e., gross domestic product, minus

⁴To be more specific, the authors identify six banking policy measures (such as nationalizations) and require at least three of these measures to have been implemented in order to consider a crisis as systemic. Other criteria that are taken into account are discussed in [Laeven and Valencia \(2020\)](#).

consumption of fixed capital, plus net foreign income). To the best of our knowledge, the WID is the only database offering systematic and comparable measures of income deciles.

The pretax income distribution indicators used in this study cover (i) top incomes, (ii) the middle-class and (iii) bottom-income groups. Top-income shares consist of the pre-tax national income shares held by the top 1%, top 10% and top 20% richest. Different top-income shares are considered because they feature important heterogeneity: labor incomes constitute the bulk of top deciles, while the 1% richest onwards rely more on capital and business incomes (see [Roine and Waldenström \(2015\)](#), for instance). The middle class is measured by the national income share held by households positioned between percentiles P21 and P79, thereby including both the lower and upper middle classes. Finally, bottom-income groups include the shares of national income held by the 20% and 10% poorest. Table [A2](#) in the Appendix presents some summary statistics for the data on income distribution.

Other variables

The determinants of income inequality that constantly appear in the literature include globalization, financial development, and public spending as well as institutions and political factors. Most of the control variables used here are obtained from the World Development Indicators (WDI) database of the World Bank.

Regarding globalization, we use the ratio of trade (i.e., the sum of exports and imports) to GDP along with the well-known [Chinn and Ito \(2006\)](#)'s index measuring a country's capital mobility status. Using a panel of 51 countries over the 1981–2003 period, [Jaumotte et al. \(2013\)](#) find that "*whereas trade globalization is associated with a reduction in inequality, financial globalization—and foreign direct investment in particular—is associated with an increase in inequality*". Furthermore, while the role of financial development—measured in our paper by the ratio of domestic private credit to GDP—has been repeatedly pointed out in the inequality literature, its net effect on income distribution is still debated; financial development could make access to credit easier for low-income households, but growing evidence shows that more finance favors top incomes and exacerbates macroeconomic volatility (see [De Haan and Sturm \(2017\)](#) and [Phelan \(2016\)](#), for instance). Political institutions are assessed by the Polity2 index, which scales the regime in place from -10 (hereditary monarchy) to +10 (consolidated democracy), and public spending is proxied by the ratio of government expenditures to GDP. Finally, we include the country's GDP and its squared term to capture the confounding effects of the Kuznets curve.

3 Prima facie evidence

To obtain a preliminary picture of the relationship between banking crises and income distribution at the national level, we estimate a standard dynamic panel model. Different income indicators are regressed on their lagged value, a banking crisis dummy, taking the value one in case of systemic banking crisis, and a set of country and time fixed effects. The results are reported in Table 1. We refrain hitherto from making any causality statement.

Columns (1) and (2) present the association between banking crises and aggregated income measured by real GDP and real national income, respectively. The trigger of a banking crisis is associated with a decrease of 3.7% in real GDP and 3.8% in national income. However, these observations conceal existing differences between income groups. Columns (3) to (8) explore the dependencies between banking crises and average income for several groups. The top 1% of earners have incomes that are 5.47% less important following a banking crisis. The estimates follow a downward trend (in absolute term) with income groups up to the middle class. The coefficient of interest is -0.044 for P10, -0.042 for P20 and -0.033 for the middle class. This pattern reverses for the 20% poorest, as the average income of B20 is reduced by 5.46% and by 8.14% for B10. This suggests that the effect of banking crises is stronger at the left and right tails of the income distribution, which justifies focusing on percentiles rather than the Gini index, for instance. In the following, we will use the income share of different percentiles to analyze the distributional effects of banking crises.

Table 1: Banking crisis and distribution of national income

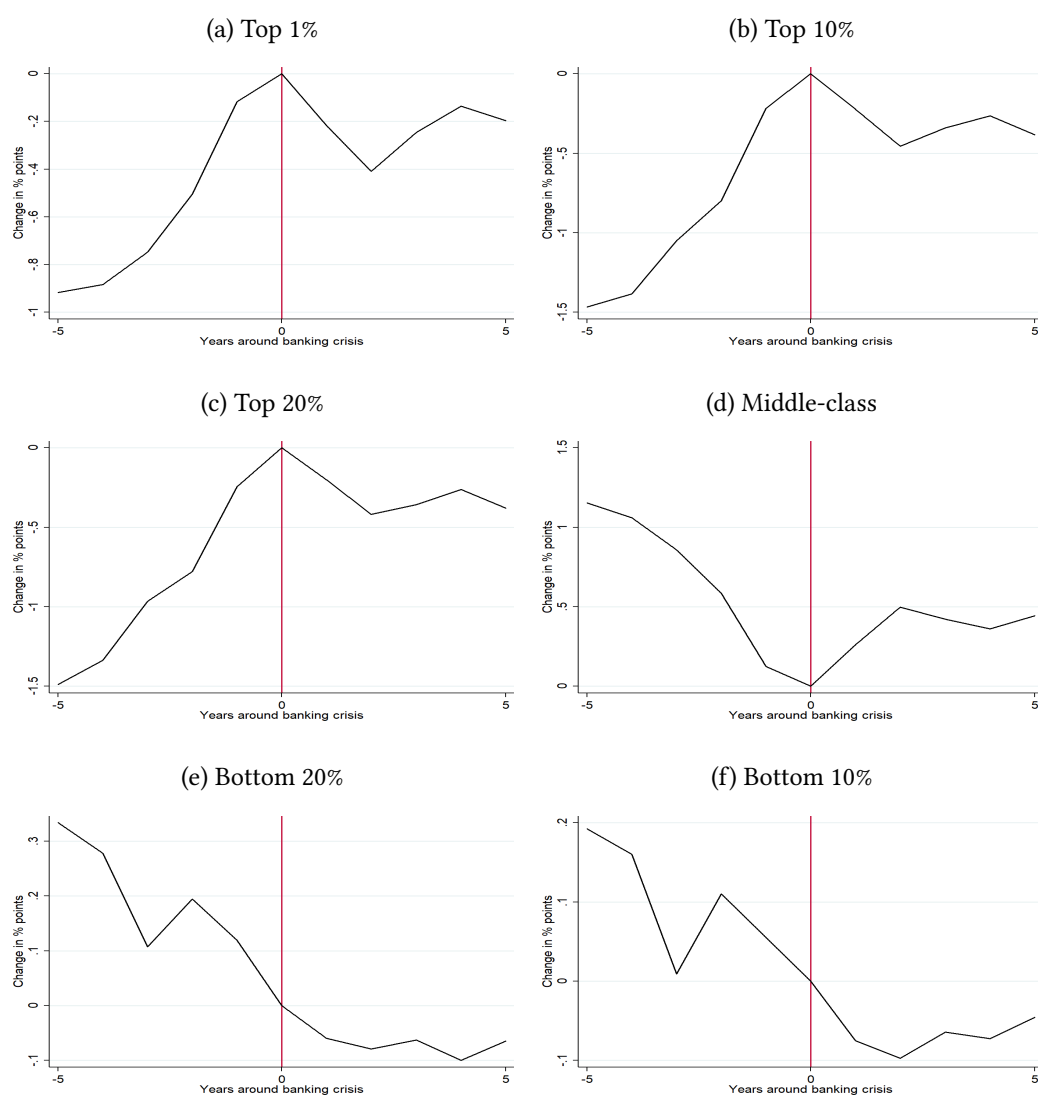
	GDPpc	NIpc	P1	P10	P20	Mid 21-79	B20	B10
Banking crisis	-0.037** (0.016)	-0.038** (0.016)	-0.054*** (0.020)	-0.044** (0.01)	-0.042** (0.017)	-0.033** (0.016)	-0.054*** (0.035)	-0.081*** (0.026)
Observations	2480	2482	2482	2471	2439	2439	2437	2428
Countries in sample	137	137	137	137	137	137	137	137

Note: This table reports prima facie evidence of the relationship between a banking crisis and national income as well as the average pre-tax income of different income groups. GDPpc and NIpc refer to the log of the real GDP per capita and the log of pre-tax national income, respectively. P1, P10, P20, Mid 21-79, B20 and B10 correspond to the log of the average pre-tax income of the top 1%, top 10% and top 20% richest, the households positioned between P21 and P79 and the bottom 20% and 10% poorest, respectively. The estimates are from dynamic panel regressions with country and time fixed effects. Cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Another way to obtain a sense of the interaction between banking crises and income distribution is to look at the dynamics of the different income percentiles around the banking crisis. Figure 1 depicts income share dynamics in countries that endure a banking crisis at year 0 relative to other countries that are spared from a massive disruption to their banking system. It is interesting to note that banking crises are preceded by different variations across income percentiles.

The shares of the top 1%, top 10% and top 20% increase prior to a banking crisis compared to countries that do not experience one, which is consistent with the observation of [Morelli \(2018\)](#) in the U.S. context. Figure (a) can be read as follows: 5 years before the trigger of a banking crisis, the income share of the top 1% is 1 percentage point lower than its value at the beginning of the crisis relative to countries that do not enter in crisis. This shows that the pre-event trend is not the same. The increase in top-income shares obviously induces income losses for other groups: figures (d), (e) and (f) show that the middle-class and bottom groups experience a reduction in their national income shares compared to countries that do not experience a banking crisis.

Figure 1: Change in income share around banking crisis



Note: Income shares before and after a banking crisis. Income shares are normalized relative to the income share prevailing the year before the banking crisis.

This observation motivates the necessity to consider a thorough empirical strategy analyzing the causal distributional effects of banking crises. In fact, the difference in pre-event trends violates the assumptions underlying panel data estimates or the difference-in-difference framework, i.e., the parallel-trends assumption. We need to explicitly model the pretrend to make it parallel or, in other words, to remove the confounding effects of income share dynamics on banking crises. This implies considering a specification that includes one or more autoregressive terms. In doing so, we address the issue of potential endogenous selection into banking crises and ensure that the error term of the model we estimate is serially uncorrelated, given that income shares show persistent serial correlation.

However, introducing autoregressive terms is not enough to fully control for all the sources of endogeneity, as there are factors other than inequality to explain the selection into a banking crisis. For instance, [Borio and Drehmann \(2009\)](#) argue that strong increases in credit and asset prices have tended to precede banking crises. As shown in [El Herradi and Leroy \(2020\)](#), changes in the income distribution may also have a nonnegligible impact on credit expansion. Furthermore, omitted variable bias can be caused by time-varying unobservable factors affecting both the likelihood of a banking crisis and the income distribution. In summary, there are many challenges to properly inferring causal assessments of banking crises on the distribution of income. We discuss our proposals to tackle these challenges in the next section.

4 Empirical approach

The design of our empirical strategy is threefold and broadly follows [Acemoglu et al. \(2019\)](#). We start with a panel model that controls for the influence of the lagged terms of the income shares as well as other potential confounding factors. Then, we propose an IV strategy to address potential endogeneity bias. Finally, we confront the endogeneity of selection by modeling through observables the selection of countries into banking crises.

Dynamic panel model

Our first empirical specification consists of estimating the following dynamic panel:

$$y_{p,i,t} = \beta Crisis_{i,t} + \sum_{j=1}^l \lambda_j y_{i,t-j} + \kappa x_{i,t-1} + \alpha_i + \gamma_t + \varepsilon_{i,t} \quad (1)$$

where $y_{p,i,t}$ refers to the income share of the percentile p in country i at time t . $Crisis$ is a dummy variable taking the value of one if country i is facing a systemic banking crisis in year t according to the crisis dates of [Laeven and Valencia \(2020\)](#). $x_{i,t-1}$ is a vector of lagged control variables, while α_i , γ_t and $\varepsilon_{i,t}$ refer to country fixed effects, time fixed effects and an error term,

respectively. The number of lags l of the dependent variable is set to 4 to eliminate residual serial correlation in the error term.

The baseline model is estimated by a fixed effect estimator. We acknowledge that the estimates are subject to an asymptotic bias because demeaning in a dynamic panel model results in correlation between the error terms and regressors, i.e., Nickell bias. However, the size of this bias decreases as the length of the sample increases. Hence, its importance in our analysis is small, as the time dimension of our panel is large, covering a maximum of 43 years and an average of almost 19 years. To ensure that our results are not affected by such a bias, for robustness we use the GMM Arellano-Bond estimator. While the latter deals with the fixed effects bias, it also introduces another one due to the proliferation of instruments ("too many instruments problem" (Roodman, 2009)) stemming from the large time dimension of our panel. The trade-off between the OLS and GMM biases we face leads us to prefer the simple OLS estimation approach as the baseline.

To make causal assessment from equation 1, we must assume that banking crises are orthogonal to contemporaneous shocks in income shares. This exogeneity assumption is strong, but it is not implausible because we control for relevant factors that simultaneously affect income shares and the experience of a banking crisis. Our specification indeed accounts for the fact that banking crises (i) are preceded by different dynamics across the income distribution and (ii) could depend on different levels of financial development, GDP or trade openness, for instance. However, a time-varying omitted variable, that is, the joint determination of income distribution and banking crises, may still confound our estimates. For this reason, we relax the exogeneity assumption of the banking crisis and propose an IV strategy.

Instrumental variable approach

Identifying causality through instrumental variables is often challenging, as it implies finding exogenous perturbations that affect the probability of a country experiencing a systemic banking crisis without directly influencing income distribution. To meet the IV conditions (relevance and exclusion restriction), we propose instruments that exploit the geographically diffusive character of banking crises. Our instruments are inspired by Acemoglu et al. (2019) and Lang and Tavares (2018) in that they use geographic transmission as a way to instrument democracy and globalization, respectively. The authors assume that a country's political regime or degree of openness is exogenously affected by the political regime or globalization intensity of the neighboring countries. We replicate this idea in our context and assume that the past manifestation of banking crises near a country affects its probability of experiencing a banking crisis. The fact

that banking crises spread geographically and produce contagion effects is well established in the literature (Cetorelli and Goldberg, 2011; Dungey and Gajurel, 2015; Kaminsky and Reinhart, 2000).

The exclusion restriction of our IV strategy imposes that banking crises abroad do not affect domestic income distribution through channels other than the (predicted) country’s banking crisis, conditional on the included control variables and fixed effects. This is questionable as long as banking crises abroad may directly reduce national income through a reduction in foreign incomes (which are considered in the WID definition of national income). The main issue is that foreign incomes are most likely earned by top percentile groups, meaning that banking crises abroad may directly impact home income distribution. To ensure that the identification strategy is not affected by this confounding effect, we add covariates measuring foreign GDP and foreign income distribution. By doing so, the contagion of a banking crisis becomes orthogonal to the geographical spread of national income and its distribution.

Taking geography into account to build our instrument seems both relevant and consistent with the exclusion restriction, but the challenge is still to concretely define geographic proximity. First, we follow Acemoglu et al. (2019) in defining the world’s different geographic regions. The definition of regions is based on the World Bank classification and leads to parceling the world into 7 regions.⁵ Each region defines the set of countries that can influence the probability of observing a banking crisis in a given country belonging to the subregion. As an illustration, we will consider that a banking crisis in Algeria is affected by banking crises in North Africa and the Middle East. To generalize, we posit that a banking crisis in country i is influenced by banking crises in the set of countries $I_i = j \neq i; R_i = R_j$, which includes countries j belonging to the same region R as country i . Then, we define our instrument, $Z_{i,t-1}$, as the one-period lagged jackknifed average of banking crisis in region R , which may be expressed as follows:

$$Z_{i,t-1} = \frac{1}{|I_i|} \sum_{j \in I_i} Crisis_{j,t-1} \quad (2)$$

where $|I_i|$ corresponds to the number of counties in the region R minus one.

In view of the usual IV threats to the exclusion restriction assumption, we introduce another geographical instrument based on the same idea that banking crises have a geographically diffusive character. On the one hand, this allows us to evaluate the consistency of our results with another source of exogenous geographical perturbations of the home banking crisis. On

⁵These regions are Africa, East Asia and the Pacific, Eastern Europe and Central Asia, Western Europe and other developed countries, Latin America and the Caribbean, the Middle East and North Africa, and South Asia. For robustness, we also consider another (larger) partition of the world following the United Nations classification. The results are broadly insensitive to the region partition choice.

the other hand, it addresses the limitation of the just-identified 2SLS model, where exclusion restrictions cannot be tested. In this respect, using two geographical instruments at the same time, i.e., overidentifying the model, enables us to perform a Hansen overidentification test and, therefore, to formally test the exclusion restriction.

While the first instrument considers that the contagion of banking crises is regional, the second instrument assumes that banking crises spatially spread according to the geographic distance between two countries. Specifically, we instrument the banking crisis of country i at time t with the one-period lag of the inverse of the geographic distance weighted banking crisis triggers in all other countries $j \neq i$ at time $t - 1$:

$$Z'_{i,t-1} = \frac{\sum_{j \neq i} \left(\frac{1}{\text{distance}_{i,j}} \text{Crisis}_{j,t-1} \right)}{\sum_{j \neq i} \frac{1}{\text{distance}_{i,j}}} \quad (3)$$

As a result, our 2SLS baseline model we estimate is given by:

$$y_{i,t,p} = \beta \widehat{\text{Crisis}}_{i,t} + \sum_{j=1}^l \lambda_j y_{i,t-j} + \kappa x_{i,t-1} + \alpha_i + \gamma_t + \varepsilon_{i,t} \quad (4)$$

$$\widehat{\text{Crisis}}_{i,t} = \delta Z_{i,t-1} + \lambda Z'_{i,t-1} + \sum_{j=1}^l \theta_j y_{i,t-j} + \kappa x_{i,t-1} + \mu_i + \eta_t + v_{i,t} \quad (5)$$

where μ_i , η_t and $v_{i,t}$ refer to country fixed effects, time fixed effects and the error term, respectively. Our 2SLS model extends equation 1 by strictly relaxing the exogeneity of the banking crisis variable. Note that we will also estimate a 2SLS model with one instrument at each time.

Treatment effects

Our third empirical strategy uses a treatment effect framework, in which the treatment is the trigger of a banking crisis. In comparison to the previously mobilized methods, this approach allows us to estimate the dynamic effects of a banking crisis on income distribution. The main challenge with such an approach is that banking crises are not randomly assigned. Countries that experience a banking crisis are different in terms of their potential outcomes, as highlighted by our prima facie evidence in section 3. As a consequence, this raises a causal inference problem, and the average treatment effect on the treated (ATT) cannot be directly obtained as in a randomized control trial.⁶

⁶ATT is the difference between the change of income distribution given the treatment and the change of income distribution with no treatment for a treated country.

To obtain ATT from our observational study, we must isolate the effects of the banking crisis on income distribution from other observable factors affecting both the treatment assignment and the outcomes. Accounting for the confounding effects of the covariates reduces the selection bias (i.e., omitted variable bias), which allows us to properly compare the outcome between the treatment and control groups. The general approach consists of estimating a propensity score defined as the likelihood of treatment assignment conditional on observed covariates. Formally, the ATT is given by:

$$ATT = E_{e(X)}(E(Y_1|e(X), Z = 1) - E(Y_0|e(X), Z = 1)) \quad (6)$$

where $e(X)$ is the propensity score to experience a banking crisis. X refers to a vector of covariates, Y_1 indicates the potential outcome when the country is in the treated group, Y_0 is the potential outcome when the country is in the control group, and Z is a dummy indicating the treatment selection.

To estimate ATT, we mobilize the inverse probability weighting (IPW) estimator as well as the doubly robust estimator.⁷ The IPW method weighs each outcome by the inverse probability of being treated, which is obtained from a logit model. In practice, this method gives more weight to the outcome in the control group (no banking crisis) with a high propensity score to experience a banking crisis. In addition to time fixed effects, our logit model includes as covariates the one- to four-period lags of the income segments, the log of GDP and the credit-to-GDP ratio as well as a commodity term of the trade index. In fact, it has been demonstrated that credit growth is a strong predictor of banking crises on average (see [Borio et al. \(2002\)](#), [Drehmann et al. \(2011\)](#), [Jordà et al. \(2011\)](#) [Schularick and Taylor \(2012\)](#), among others), while commodity prices could help predict banking crises specifically in low-income countries ([Eberhardt and Presbitero, 2018](#)).

Our estimations of the ATT are supplemented by the doubly robust estimator. The latter complements the IPW estimator by adjusting the outcome using a linear regression model. First, we estimate the propensity score and weigh outcomes; then, we regress these weighted outcomes on covariates. The covariates we use here are the same as those mobilized in the logit model, in addition to the controls included in the two previous empirical strategies (government expenditures, trade and financial openness, political regime).

The dynamic causal effect is afterwards obtained by estimating a sequence of ATT over time, in which the endogenous variable (the national income share held by the top 1% richest, for instance) is allowed to vary. More precisely, ATT_1 will indicate the change in the considered

⁷In the Appendix, we also report ATT results from regression adjustment methods.

income share between period -1 and the starting year of the banking crisis (period 0) caused by the banking crisis, while ATT_5 will indicate the change in income share between period 5 and period -1.

As long as the one- to four-year lags of the income share are included as control variables, we ensure that there is no pre-event trend for the five years before the banking crisis. However, a pretrend could be observed even before, making the no pretrend assumption of the ATT approach invalid. That is why we complement our approach by estimating the effects of a banking crisis on backward changes in the outcome, that is, the changes in the considered income share between year -10 and year -1, for instance. This should be viewed as a simple specification test, as we do not expect significant effects of the banking crisis in period -1 on the backward dynamic changes in income distribution. In the case where the average treatment effect on the treated at period -10 (ATT_{-10}) is significantly different from 0, we should conclude that our model is not correctly specified and, therefore, projection estimates are mistaken.

5 Empirical results

In this section, we present the results of the empirical assessment of the banking crisis effects across the income distribution. The baseline and IV results are first discussed before turning to the dynamic effects of banking crises.

Dynamic panel estimates

Our first approach to estimating the distributional effects of banking crises is to introduce a full dynamic model for several indicators of income distribution. Table 2 reports the baseline results obtained from the estimation of equation 1 by a fixed effects estimator, controlling for a number of lags. The occurrence of a banking crisis is associated with significant yet differentiated effects across the income distribution. First, banking crises and top-income shares are negatively related, with the richest 1% bearing the largest losses. Columns (1)-(3) report the coefficients of the banking crisis dummy for top incomes and indicate that the shares of national income held by the top 1%, top 10% and top 20% are 0.28, 0.27 and 0.22 percentage points lower, respectively. Second, column (4) shows, by contrast, that this effect reverses for the middle class, as banking crises are associated with an increase of 0.26 percentage points in the pretax income share going to households positioned between P21 and P79.⁸ Finally, as far as bottom-income groups

⁸The effects are, however, heterogeneous within the middle class. The results reported in Table A3 of the Appendix show that the upper middle class is less penalized by a banking crisis and benefits more from the positive distributional effects following a banking crisis than the lower middle class. One possible explanation for this result is that the labor demand of the upper middle class is less sensitive to macroeconomic shocks than that of the lower middle class.

are concerned, the reported coefficients in columns (5)-(6) show a negative correlation between banking crises and the income shares of the 20% and 10% poorest. It should also be noted that low-income group losses from a banking crisis are only one-third of those experienced by top incomes, thereby suggesting that the negative effects of banking crises are stronger at the right tail of the income distribution.

As mentioned, the baseline estimations of Table 2 also control for 4 lags of the pretax national income shares. In a pattern common to all of the specifications that we report, we find a sizable amount of persistence in the distribution of income, with coefficients of the first lag being significantly positive and smaller than 1, while those of the third and fourth lags are negative, especially for top-income shares. Regarding other control variables, we find that the amount of government expenditures is positively correlated with top-income shares, but this correlation is negative for the middle class. Conversely, the enhancement of a country's institutional environment, i.e. democratic consolidation, counts against top incomes and favors middle-class households, while its effect on the bottom of the income distribution is insignificant.

The robustness of the dynamic panel estimates is assessed in Table 3. The first set of checks estimates equation 1 with country and time fixed effects while omitting the set of control variables. The results are consistent with baseline estimations, although the association between banking crises and bottom-income shares is not statistically significant. In the second battery of checks, we consider specifications with smaller lag numbers and find that the effect of interest holds only for the richest 1% and bottom-income shares, whereas the coefficient for the middle class is only significant at the 10% level. Then, we assess the sensitivity of our results to outliers by taking into account the prediction errors in the regressions. This procedure shows that top-income shares and, to a lesser extent, bottom-income groups are negatively affected by a banking crisis, while the coefficient of the middle-class turns insignificant. An additional set of sensitivity analyses includes the contemporaneous and three lags of the control variables. Although the implied dynamics are now richer, the overall effect of banking crises on the income distribution is close to that found in Table 2, except that estimates for the middle-class and poor households are less statistically significant. The last set of checks uses the Arellano-Bond GMM estimator that deals with the Nickell bias and produces estimates that do not depart from the general pattern established in the baseline findings.

Table 2: Banking crisis and the distribution of national income - Baseline results

	Top 1%	Top 10%	Top 20%	Middle-class	Bottom 20%	Bottom 10%
Banking crisis	-0.280*** (0.094)	-0.276** (0.117)	-0.218** (0.106)	0.258*** (0.090)	-0.063* (0.035)	-0.051*** (0.015)
Lag1	0.911*** (0.107)	0.9777*** (0.065)	0.988*** (0.052)	0.953*** (0.054)	0.6651*** (0.136)	0.403*** (0.150)
Lag2	0.130 (0.119)	0.090 (0.082)	0.058 (0.065)	0.067 (0.065)	0.153*** (0.034)	0.148*** (0.027)
Lag3	-0.106* (0.064)	-0.155*** (0.049)	-0.141*** (0.049)	-0.084 (0.058)	0.003 (0.052)	0.093*** (0.030)
Lag4	-0.063 (0.043)	-0.014 (0.032)	-0.006 (0.034)	-0.046 (0.038)	0.010 (0.026)	0.031 (0.042)
GDP	0.507 (0.453)	0.081 (0.446)	0.045 (0.382)	-0.051 (0.346)	0.009 (0.127)	0.002 (0.086)
GDP ²	-0.013 (0.010)	-0.001 (0.011)	-0.000 (0.009)	0.001 (0.008)	-0.001 (0.003)	-0.000 (0.002)
Gov.	0.029** (0.012)	0.029** (0.011)	0.021** (0.010)	-0.018** (0.009)	-0.005* (0.003)	-0.003 (0.002)
Trade	0.001 (0.001)	0.000 (0.002)	0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.000)
Fin. Open.	0.026 (0.147)	0.084 (0.174)	0.103 (0.150)	-0.128 (0.144)	0.031 (0.050)	0.028 (0.038)
Credit	0.002 (0.002)	0.000 (0.001)	0.001 (0.001)	-0.001 (0.001)	-0.001*** (0.000)	-0.001*** (0.000)
Polity2	-0.028** (0.012)	-0.031** (0.013)	-0.027** (0.011)	0.025** (0.010)	0.006* (0.003)	0.004 (0.003)
Observations	2634	2618	2566	2566	2566	2566
Countries in sample	140	140	140	140	140	140

Note: This table shows baseline results from the estimation of equation 1 for the national income shares held by the top 1% (column (1)), top 10% (column (2)) and top 20% richest (column (3)); the middle class (column (4)); the national income share held by the 20% and 10% poorest. Country and time fixed effects are included and cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3: Banking crisis and distribution of national income - Baseline robustness

	Top 1%	Top 10%	Top 20%	Middle-class	Bottom 20%	Bottom 10%
1- No controls	-0.170** (0.067)	-0.233*** (0.080)	-0.211*** (0.075)	0.224*** (0.061)	-0.020 (0.033)	-0.037* (0.021)
2- One lag	-0.271** (0.110)	-0.205 (0.132)	-0.143 (0.117)	0.188* (0.104)	-0.057** (0.026)	-0.038*** (0.012)
3- Two lags	-0.256** (0.107)	-0.181 (0.125)	-0.127 (0.111)	0.173* (0.098)	-0.066** (0.031)	-0.045*** (0.013)
4- Outliers	-0.266*** (0.090)	-0.278** (0.110)	-0.216** (0.099)	-0.078 (0.185)	-0.068** (0.034)	-0.054*** (0.015)
5- More controls	-0.265*** (0.098)	-0.246** (0.124)	-0.190* (0.113)	0.247* (0.132)	-0.068* (0.037)	-0.026* (0.015)
6- GMM	-0.280** (0.115)	-0.241* (0.143)	-0.182 (0.131)	0.223** (0.110)	-0.059 (0.036)	-0.036** (0.015)

Note: This table reports the relationship between banking crises and the considered income distribution measures for specifications that (i) remove the vector of control variables, (ii) consider only one/two lags, (iii) account for outliers, (iv) add more controls and (v) use the Arellano–Bond GMM estimator. Country and time fixed effects are included and cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Control variables and fixed effects are not reported.

IV estimates: banking crises waves

Our second empirical approach exploits the regional diffusion of banking crises as an exogenous source of variation in a country’s likelihood of experiencing a disruption of its banking system. Such a strategy deals with time-varying omitted variables that simultaneously affect banking crises and income distribution, which are not addressed by the standard dynamic panel model. To the best of our knowledge, this is the first contribution attempting to examine the causal effect of banking crises on different segments of the income distribution.

Table 4 presents our 2SLS estimates of equation 5. These estimates depart from the baseline results in two important respects: (i) the (causal) effect of a banking crisis on top incomes and the middle class is larger than in previous estimates, whereas (ii) its impact on bottom-income groups is positive but not statistically significant. Such a discrepancy with the previous findings speaks to the existence of an attenuation bias in the OLS estimation. To put it differently, our IV captures waves in banking crises that are more likely to reflect exogenous fluctuations than those we could identify using a simple OLS regression control strategy. In particular, the banking crisis coefficient for the top 20% richest is -0.93 percentage points (column (3)), while that of the middle class reports an increase of 0.91 percentage points (column (4)). It therefore seems that banking crises redistribute income from rich households to the middle class without affecting bottom-income groups. This redistribution is consistent with [Bazillier and Najman \(2017\)](#), suggesting that banking crises primarily affect capital incomes, which constitute the main earnings source

of well-off households.

The inclusion of two instruments in equation 5 further enables us to perform a Hansen over-identification test, which provides no evidence of misspecification. The Kleibergen-Paap test F-statistics are also reported and reject the null hypothesis of weak instruments. In addition, the overall amount of persistence of income distribution indicators, reported below the banking crisis coefficients of Table 4, is close to that found in the dynamic panel estimates.

We carry out a number of tests to check the robustness of our IV findings. These are shown in Table 5. Because we adopted an overidentified specification with two instruments, a first test consists of assessing the sensitivity of our results to only one instrument. First, exploiting the exogenous diffusion of banking crises through regional waves confirms the results reported before: top-income shares decline in the aftermath of a banking crisis and the income share of the middle-class increases, while low-income households are unaffected. Second, assuming that banking crises spread spatially according to geographical distance suggests that top deciles and the middle class follow the aforementioned pattern, while the estimated coefficient on the richest 1% is negative but not statistically significant. Further, to assess whether the IV exclusion restriction is violated, we estimate equation 5 without control variables and show that the baseline IV findings continue to hold. Another potential source of bias in our IV estimates could arise from differential trends in the distribution of income among countries that have experienced a banking crisis. For this reason, we control for these trends by including the average, per region, of top-income shares and middle-class and bottom-income shares. The results are reported in line (4) of Table 5 and show strong consistency with the main IV estimates.

Table 4: Banking crisis and the distribution of national income - IV results

	Top 1%	Top 10%	Top 20%	Middle-class	Bottom 20%	Bottom 10%
Banking crisis	-0.776** (0.332)	-1.081*** (0.330)	-0.929*** (0.287)	0.906*** (0.257)	0.209 (0.150)	0.127 (0.114)
Lag1	0.906*** (0.109)	0.978*** (0.068)	0.991*** (0.055)	0.950*** (0.057)	0.663*** (0.137)	0.406*** (0.156)
Lag2	0.144 (0.120)	0.111 (0.085)	0.076 (0.068)	0.088 (0.067)	0.151*** (0.032)	0.145*** (0.034)
Lag3	-0.112* (0.064)	-0.181*** (0.049)	-0.167*** (0.050)	-0.099 (0.062)	-0.015 (0.045)	0.079*** (0.025)
Lag4	-0.065 (0.044)	-0.003 (0.032)	0.005 (0.033)	-0.041 (0.041)	0.015 (0.032)	0.031 (0.051)
GDP	0.758 (0.606)	-0.156 (0.626)	-0.149 (0.567)	0.092 (0.499)	0.136 (0.199)	0.101 (0.132)
GDP ²	-0.019 (0.015)	0.005 (0.016)	0.005 (0.015)	-0.003 (0.013)	-0.004 (0.005)	-0.003 (0.003)
Gov	0.031*** (0.012)	0.031*** (0.011)	0.021** (0.009)	-0.019** (0.009)	-0.005 (0.003)	-0.003 (0.002)
Trade	0.001 (0.001)	-0.000 (0.002)	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Fin. Open.	-0.005 (0.153)	0.070 (0.174)	0.103 (0.148)	-0.124 (0.143)	0.037 (0.057)	0.029 (0.041)
Credit	0.004** (0.002)	0.003* (0.002)	0.004** (0.002)	-0.003* (0.002)	-0.002** (0.001)	-0.002** (0.001)
Polity2	-0.030** (0.012)	-0.033** (0.013)	-0.028** (0.012)	0.026** (0.010)	0.007* (0.004)	0.004 (0.003)
Observations	2555	2539	2487	2487	2487	2487
Countries in sample	137	137	137	137	137	137
Kleibergen-Paap	25.0	24.1	23.6	23.6	25.8	24.9
Hansen <i>p</i> – value	0.37	0.92	0.77	0.72	0.71	0.62

Note: This table shows the IV results from the estimation of equation 5 for the national income shares held by the top 1% (column (1)), top 10% (column (2)) and top 20% richest (column (3)); the middle class (column (4)); the national income share held by the 20% and 10% poorest. We report the Kleibergen and Paap (2006) statistic for weak instruments and the p-value for the Hansen over-identification test. Country and time fixed effects are included and cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Similarly, to ensure that the diffusion of banking crises is not caused by other regional trends, we include region-specific trends in our specification. The estimations displayed in line (5) suggest that our IV results are not affected by unobserved regional heterogeneity. In dealing with regionally correlated omitted variables, two complementary procedures are introduced, which consist of controlling for foreign income distribution and foreign GDP. First, income distribution indicators are allowed to be spatially correlated as a function of the inverse of the distance between countries (see equation 3). The results reported in line (6) are consistent with our baseline findings. Second, we consider a specification that includes GDP per capita at the regional level as well as countries' "spatial" GDP, which is a function of the geographical distance between two countries belonging to the same region. The results shown in line (7) remain similar to the baseline evidence. Finally, we explore the sensitivity of our IV results to outliers. As we did in the previous dynamic panel model and following [Acemoglu et al. \(2019\)](#), our IV specification is estimated by excluding countries with a standardized residual above 1.96 or below -1.96 . The evidence reported in line (8) of Table 5 is very similar to our baseline results, suggesting that our findings are not driven by outliers.⁹

Table 5: Banking crisis and the distribution of national income - IV robustness

	Top 1%	Top 10%	Top 20%	Middle-class	Bottom 20%	Bottom 10%
1- Regional instrument	-0.989** (0.476)	-1.067** (0.523)	-0.823** (0.413)	0.770** (0.387)	0.235 (0.202)	0.149 (0.152)
2- Spatial instrument	-0.477 (0.442)	-1.037** (0.482)	-1.029** (0.431)	1.032** (0.413)	0.170 (0.134)	0.092 (0.099)
3- No controls	-0.711** (0.320)	-1.064*** (0.334)	-0.908*** (0.293)	0.832*** (0.247)	0.148 (0.106)	0.059 (0.072)
4- Regional income	-0.745** (0.328)	-1.019*** (0.324)	-0.890*** (0.287)	0.867*** (0.258)	0.201 (0.148)	0.135 (0.120)
5- Regional trend	-0.712** (0.336)	-1.071*** (0.348)	-0.905*** (0.305)	0.816*** (0.255)	0.165 (0.110)	0.076 (0.073)
6- Spatial income dist.	-0.790** (0.332)	-1.067*** (0.332)	-0.920*** (0.288)	0.898*** (0.258)	0.214 (0.153)	0.128 (0.115)
7- Spatial & regional GDP	-0.704** (0.321)	-0.983*** (0.319)	-0.858*** (0.281)	0.842*** (0.250)	0.193 (0.148)	0.124 (0.115)
8- Outliers	-0.776** (0.332)	-1.081*** (0.330)	-0.929*** (0.287)	1.272*** (0.360)	0.209 (0.150)	0.127 (0.114)

Note: This table performs several robustness checks on estimates from the IV approach. Equation 5 is estimated (i) using the regional instrument, (ii) the spatial instrument; (iii) without control variables; (iv) accounting for regional income shares trends, (v) unobserved regional trends, (vi) spatial income distribution effects; (vii) spatial & regional GDP outliers and (viii) outliers. Country and time fixed effects are included and cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

⁹Table A4 also reports the effect of banking crises on various income percentile ratios.

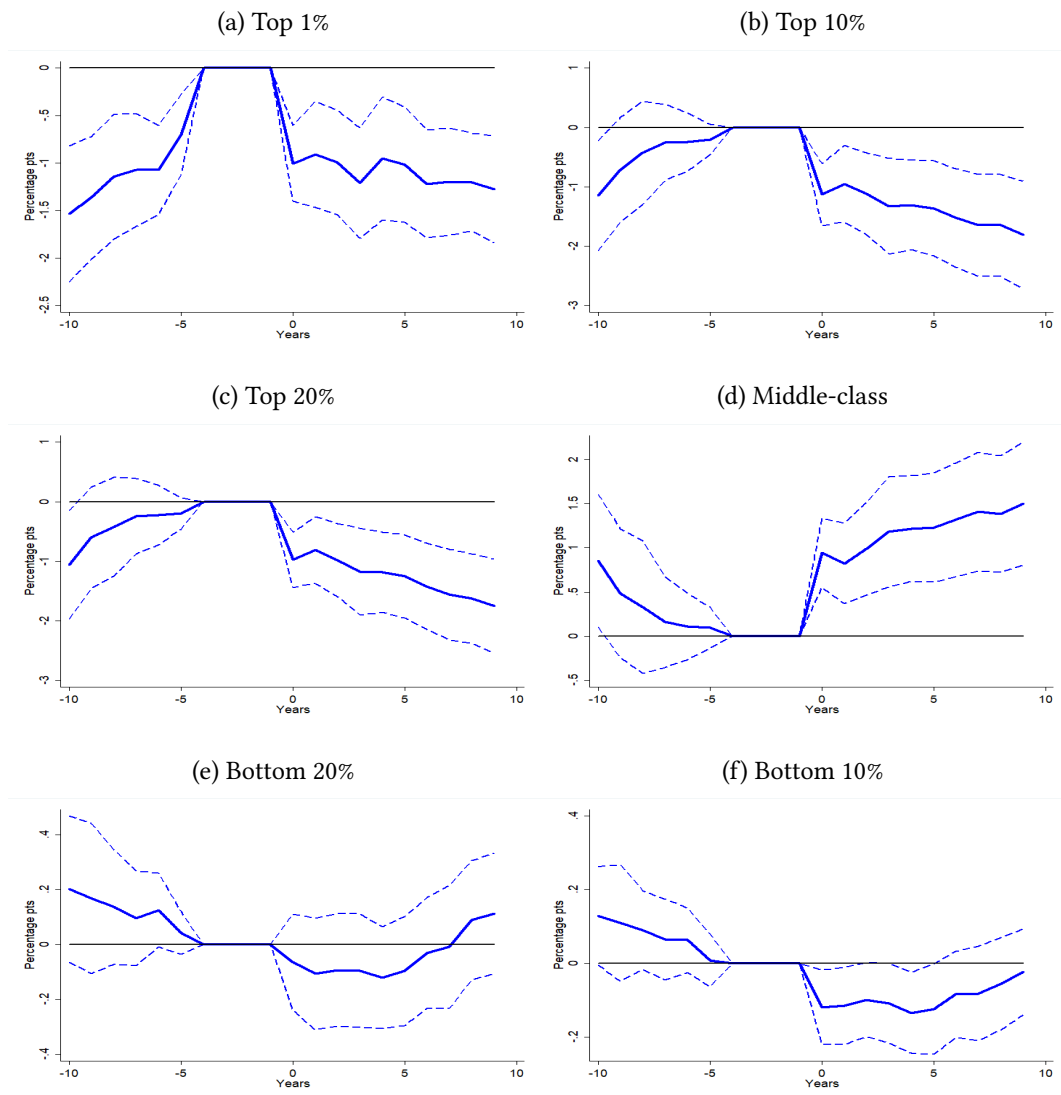
5.1 Treatment effects and semiparametric estimates

The previous subsection allowed us to draw causal assessment while controlling for income distribution dynamics. Although our IV approach produces consistent estimates under its maintained assumptions, it embodies two fundamental limitations: (i) the relationship between the income segments and banking crises is assumed to be linear and (ii) the time pattern of the cumulative effects of interest is restricted.

To address these shortcomings, we estimate the dynamic effects of a banking crisis on the subsequent path of income distribution indicators by modeling the selection of countries into banking crises. In practice, two different empirical methodologies are mobilized. First, we follow [Angrist and Guido \(2011\)](#) and [Angrist et al. \(2018\)](#) in estimating the effect of banking crises on income distribution, conditioned on the propensity score for the occurrence of a banking crisis. Specifically, this propensity score is estimated using a logit regression of the frequency of banking crises on year fixed effects and four lags of the income segment considered, the log of GDP and the credit-to-GDP ratio. [Figure 2](#) depicts, for each income segment, the estimates in the periods $p = -10, -5, \dots, 10$ with $p = 0$ corresponding to the year a banking crisis has occurred. As discussed in subsection 4, the estimates for negative values of p only serve as a specification test and should not be affected by the occurrence of banking crises. The solid line plots the dynamic estimated effects of a banking crisis on different income segments (in percentage points), and the dotted lines plot its 90 percent confidence interval.

The results in [Figure 2](#) confirm the pattern described in our IV findings and show that there are no differential trends in income segments before the trigger of a banking crisis. Top incomes experience a sharp decline in pretax national income shares (graphs (a), (b) and (c)), while the share of middle-class households increases systematically (graph (d)). For instance, following a banking crisis, the top 10% income share declines by 1.16 percentage points and reaches a negative impact of -2.06 percentage points in year 10 after the crisis. Conversely, the income share going to the middle class increases by 0.97 percentage points in the first year and stands at 1.97 percentage points ten years later. It is important to note that this does not necessarily mean that middle-class households are better off after a banking crisis. Rather, it suggests that the share of national income held by the middle class has partially recovered from the downward trend it experienced before the onset of the banking crisis. As far as bottom-income groups are concerned, the effects are not clear-cut; banking crises negatively affect low-income households, but this impact is not statistically significant, especially for the 20% poorest (see graphs (e) and (f) of [Figure 2](#)).

Figure 2: ATET based on the IPW estimator



Note: The figures show the over-time average treatment effect on the treated (ATET) of banking crises on the income share of different income groups. The ATET are obtained from the inverse probability weighting approach and are depicted by solid lines with 90% confidence bands based on cluster-robust standard error estimates in dashed lines.

Second, we use a "doubly robust estimator", which combines a linear regression of changes in the income distribution p periods after a banking crisis with the propensity score of the latter, which is estimated from a logit model. In [Acemoglu et al. \(2019\)](#) terms, such an estimator "*both reweights observations in the control group by their propensity score and adjusts the counterfactual outcome using a linear regression model*". This "double" method therefore improves the reliability of our estimations because, as explained by [Imbens and Wooldridge \(2009\)](#), it is consistent if either the linear model for potential outcomes or the logit model for banking crises is valid. In addition to the lags of the income segment, log of GDP, credit-to-GDP ratio and a commodity term of the trade index, the covariates of the linear model also include government spending, trade openness, capital mobility and institutional quality.

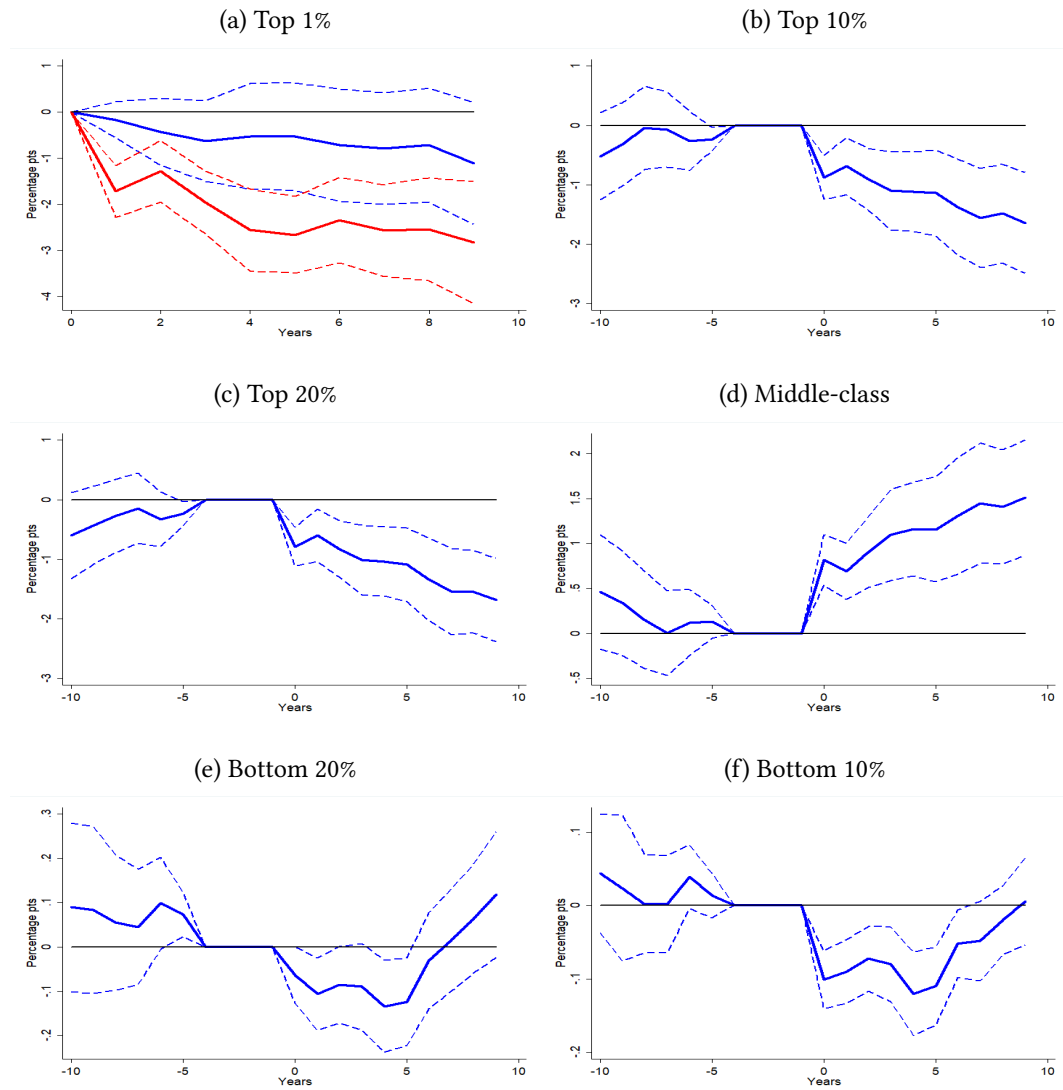
The semiparametric estimates of the dynamic effects of banking crises on the income distribution are shown in [Figure 3](#). Reassuringly, there are no trends preceding the trigger of a banking crisis, and the estimated ATET for different income segments are very similar to those from the propensity score approach. Rich households are the main losers; the immediate negative effect of a banking crisis on top-income shares ranges between -0.87 and -0.80 in percentage points for the top 1%, top 10% and top 20%, while the impacts over a ten-year horizon range between -0.89 and -1.80. The middle class recovers from the precrisis trend as its income share increases by 1.50 percentage points ten years after the crisis.¹⁰ Contrary to our previous IPW estimates, the doubly robust estimates produce significant results regarding bottom-income households. We find a significant negative effect of the treatment on the 20% and 10% poorest. This means that the income shock for the poorest households is more severe than it is for the middle class, thereby confirming our *prima facie* evidence. The dualism of the labor market can constitute a credible explanation of this result. As a matter of fact, the middle class, which mainly participates in the primary labor market, is more insulated from the uncertainty and variability of labor demand than bottom-income households.¹¹ Interestingly, we note that these effects are only transitory, as the significant impact on the 20% and 10% poorest fades over the medium run. This contrasts with the dynamic effects observed for top incomes and, most importantly, highlights that the recovery of the labor market from a banking crisis tends to be faster than that observed in financial markets, which is consistent with the fact that the financial cycle is much longer than the standard business cycle.¹²

¹⁰The results obtained using a linear regression adjustment estimator are depicted in [Figure A2](#) of the Appendix and confirm the impacts of banking crises on top incomes and the middle class.

¹¹Actually, the labor force of the latter may be viewed, to some extent, as the "residual" production factor that allows firms to respond to the productive activity fluctuations caused by a banking crisis.

¹²Note that our semi-parametric estimate strategy also allows us to test the heterogeneity of our results and give insight on the underlying mechanism of such a redistribution. In [figure A3](#), we present dynamic estimates of P20 (P20 only to save space) for the usual income classification of our sample (low income, middle income, high income) as well as for a high and low levels of: economic development (real GDP), political institution (Polity2),

Figure 3: ATET based on the doubly robust estimator



Note: The figures show the over-time average treatment effects on the treated (ATET) of banking crises on the income share of different income groups. The ATET are obtained from the doubly robust estimator and are depicted by solid lines with 90% confidence bands based on cluster-robust standard error estimates in dashed lines.

financial openness, market returns and the output gap measured just before the crisis. To obtain two regimes, we interact our banking crisis variable with a dummy variable taking the value one when real GDP is over (under) the median of the empirical distribution and zero otherwise, which gives the effect of banking crisis on income distribution when economic development is high (low). As can be seen, our results are unconditional to the level of economic and political development. In contrast the level of financial openness matters. Further, we note that the redistributive effects of a banking crisis are more severe when the financial markets are severely impacted by the banking crisis. Finally, we observe that a positive output gap before the crisis induces more loss of income for the top-income in the wake of the crisis.

6 Conclusion

This paper sought to analyze the distributional consequences of banking crises in 140 countries over the 1970-2017 period. The existing literature on the interaction between banking crises and income inequality lacks strong causal assessment and does not take interest in uncovering the differentiated effects that a banking crisis may have on segments of the income distribution. We take a step in this direction by examining the causal effect of banking crises on rich, middle-class and poor households. To do so, we combine [Laeven and Valencia \(2020\)](#)'s database on the timing of banking crises with the pretax national income shares held by the rich (top 1%, top 10% and top 20%), the middle class (households positioned between P21-P79) and low-income households (bottom 20% and bottom 10%). These data are obtained at the country level from the World Inequality Database (WID), which is one of the few databases offering systematic and comparable measures of income deciles.

Estimating the causal effect of banking crises on income distribution faces several challenges, and there is no perfect strategy to address them. Most importantly, our preliminary descriptive analysis indicates that income inequality increases in periods preceding the trigger of a banking crisis, which invalidates the parallel-trends assumption that underlies standard panel data models. Our approach broadly follows [Acemoglu et al. \(2019\)](#) in using a number of different strategies, which, reassuringly, all yield similar results.

First, we estimate a dynamic linear panel model, which includes both country fixed effects and the lags of the income shares. This strategy leads to estimates indicating that a banking crisis is negatively associated with the right and left tails of the income distribution (top- and bottom-income shares), while this correlation is positive for the middle class.

Second, we address the issue of time-varying omitted variables by using an instrumental variable approach. Specifically, we build on an extensive body of the literature showing that banking crises spread geographically and produce contagion effects across borders. The idea is to identify regional waves of banking crises that are exogenous to domestic economic conditions. Our IV estimates suggest that the effect of banking crises on top incomes and the middle class continues to hold, while that on bottom-income shares is not statistically significant.

Third, we introduce a treatment approach framework to estimate the dynamic effects of banking crises on income distribution. Such an approach models the propensity to experience a banking crisis by using lags of GDP and other covariates. The semiparametric estimators lead to fairly similar estimates; banking crises have persistent effects on top incomes and the middle class, while their interaction with bottom-income shares is not clear cut.

Our results have important implications for the design of public policies aimed at limiting the distributional consequences of banking crises. They can also be useful to policymakers to better identify the winners and losers of these crises. For future research, more steps should be undertaken to fully grasp the mechanisms underlying these effects.

References

- Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.
- Alveredo, F., Atkinson, A. B., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2016). Distributional national accounts guidelines: Methods and concepts used in wid.world. *WID.world WPS N° 2016/2*.
- Angrist, J. D. and Guido, K. M. (2011). Causal effects of monetary shocks: Semi-parametric conditional independence tests with a multinomial propensity score. *The Review of Economic Statistics*, 93(3):725–747.
- Angrist, J. D., Jordà, s., and Kuersteiner, G. (2018). Semiparametric estimates of monetary policy effects: String theory revisited. *Journal of Business & Economic Statistics*, 36(3):371–387.
- Atkinson, A. and Morelli, S. (2015). Inequality and crises revisited. *Economia Politica*, 32(1):31–51.
- Atkinson, A. B. and Morelli, S. (2011). Income inequality and banking crisis: a first look. *Economic Systems*, UNDP-HDRO Occasional Papers No. 2011/6.
- Bazillier, R. and Hericourt, J. (2017). The circular relationship between inequality, leverage, and financial crises. *Journal of Economic Surveys*, 31(2):463–496.
- Bazillier, R. and Najman, B. (2017). Labour and financial crises: Is labour paying the price of the crisis? *Comparative Economic Studies*, 55(1):55–76.
- Bellettini, G., Delbono, F., Karlström, P., and Pastorello, S. (2019). Income inequality and banking crises: Testing the level hypothesis directly. *Journal of Macroeconomics*, 62(C):103054.
- Bordo, M. D. and Meissner, C. M. (2012). Does inequality lead to a financial crisis? *Journal of International Money and Finance*, 31(8):2147–2161.
- Borio, C., Lowe, P., et al. (2002). Assessing the risk of banking crises. *BIS Quarterly Review*, December.
- Borio, C. E. and Drehmann, M. (2009). Assessing the risk of banking crises—revisited. *BIS Quarterly Review*, March.
- Cetorelli, N. and Goldberg, L. S. (2011). Global banks and international shock transmission: Evidence from the crisis. *IMF Economic review*, 59(1):41–76.
- Chaudron, R. and de Haan, J. (2014). Dating banking crises using incidence and size of bank failures: Four crises reconsidered. *Journal of Financial Stability*, 15(C):63–75.
- Chinn, M. and Ito, H. (2006). What matters for financial development? capital controls, institutions, and interactions. *Journal of Development Economics*, 81(1):163–192.
- Cho, Y. and Newhouse, D. (2013). How did the great recession affect different types of workers? evidence from 17 middle-income countries. *World Development*, 41(C):31–50.
- Chwieroth, J. and Walter, A. (2017). Banking crises and politics: a long run perspective. *International Affairs*, 93(5):1107–1129.

- De Haan, J. and Sturm, J.-E. (2017). Finance and income inequality: A review and new evidence. *European Journal of Political Economy*, 50(December):171–195.
- Denk, O. and Cournède, B. (2015). Finance and income inequality in oecd countries. *OECD Economics Department WP1224*.
- Drehmann, M., Borio, C., and Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. *International Journal of Central Banking*, 7(4):189–240.
- Dungey, M. and Gajurel, D. (2015). Contagion and banking crisis–international evidence for 2007–2009. *Journal of Banking & Finance*, 60(C):271–283.
- Eberhardt, M. M. and Presbitero, A. (2018). *Commodity Price Movements and Banking Crises*. International Monetary Fund.
- El Herradi, M. and Leroy, A. (2020). Do rising top incomes fuel credit expansion? *Economics Letters*, 196(C):109539.
- Imbens, G. and Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Business & Economic Statistics*, 47(1):5–86.
- Jaumotte, F., Lall, S., and Papageorgiou, C. (2013). Rising income inequality: Technology, or trade and financial globalization? *IMF Economic Review*, 61(2):271–309.
- Jordà, Ò., Schularick, M., and Taylor, A. M. (2011). Financial crises, credit booms, and external imbalances: 140 years of lessons. *IMF Economic Review*, 59(2):340–378.
- Kaminsky, G. L. and Reinhart, C. M. (2000). On crises, contagion, and confusion. *Journal of international Economics*, 51(1):145–168.
- Kirschenmann, K., Malinen, T., and Nyberg, H. (2016). The risk of financial crises: Is there a role for income inequality? *Journal of International Money and Finance*, 68(C):161–180.
- Laeven, L. and Valencia, F. (2020). Systemic banking crises database ii. *IMF Economic Review*, 68(2):307–361.
- Lang, V. F. and Tavares, M. M. (2018). The distribution of gains from globalization.
- Loayza, N., Ouazad, A., and Ranciere, R. (2018). Chapter 10: Financial development, growth, and crisis: Is there a trade-off? In *Handbook of Finance and Development*, page 289–311. Edward Elgar Publishing.
- Morelli, S. (2018). Banking crises in the us: the response of top income shares in a historical perspective. *Journal of Economic Inequality*, 16(2):257–294.
- Phelan, G. (2016). Financial intermediation, leverage, and macroeconomic instability. *American Economic Journal: Macroeconomics*, 8(4):199–224.
- Reinhart, C. M. and Rogoff, K. (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.
- Reinhart, C. M. and Rogoff, K. S. (2013). Banking crises: An equal opportunity menace. *Journal of Banking & Finance*, 37(11):4557–4573.

- Rhee, D.-E. and Kim, H. (2018). Does income inequality lead to banking crises in developing countries? empirical evidence from cross-country panel data. *Economic Systems*, 42(2):206–218.
- Roine, J., Vlachos, J., and Waldenström, D. (2009). The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics*, 93(7-8):974–988.
- Roine, J. and Waldenström, D. (2015). Long-run trends in the distribution of income and wealth. In *Handbook of Income Distribution*, volume 2, pages 469–592. Elsevier.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics*, 71(1):135–158.
- Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102(2):1029–61.
- van Rijckeghem, C. and Weder, B. (2001). Sources of contagion: is it finance or trade? *Journal of International Economics*, 54(2):293–308.
- van Treeck, T. (2014). Did inequality cause the u.s. financial crisis? *Journal of Economic Surveys*, 28(3):421–448.

Appendix

Table A1: Countries included in the sample

Afghanistan	Colombia	Guyana	Malta
Albania	Comoros	Haiti	Mauritania
Algeria	Congo Dem. Rep.	Honduras	Mauritius
Angola	Congo Rep.	Hungary	Mexico
Antigua and Barbuda	Costa Rica	Iceland	Moldova
Argentina	Cote d'Ivoire	India	Mongolia
Armenia	Croatia	Indonesia	Morocco
Australia	Cuba	Iran	Mozambique
Austria	Cyprus	Iraq	Myanmar
Azerbaijan	Czech Republic	Ireland	Namibia
Bahamas	Denmark	Israel	Nepal
Bahrain	Djibouti	Italy	Netherlands
Bangladesh	Dominica	Jamaica	New Zealand
Barbados	Dominican Republic	Japan	Nicaragua
Belarus	Ecuador	Jordan	Niger
Belgium	Egypt	Kazakhstan	Nigeria
Belize	El Salvador	Kenya	North Macedonia
Benin	Equatorial Guinea	Kiribati	Norway
Bhutan	Eritrea	North Korea	Oman
Bolivia	Estonia	South Korea	Pakistan
Bosnia and Herzegovina	Eswatini	Kuwait	Palau
Botswana	Ethiopia	Kyrgyz Republic	Panama
Brazil	Fij	Lao	Papua New Guinea
Brunei	Finland	Latvia	Paraguay
Bulgaria	France	Lebanon	Peru
Burkina Faso	Gabon	Lesotho	Philippines
Burundi	Gambia	Liberia	Poland
Cabo Verde	Georgia	Libya	Portugal
Cambodia	Germany	Lithuania	Qatar
Cameroon	Ghana	Luxembourg	Romania
Canada	Greece	Grenada	Russian Federation
Central African Republic	Madagascar	Malawi	Rwanda
Chad	Guatemala	Malaysia	Samoa
Chile	Guinea	Maldives	Sao Tome and Principe
China	Guinea-Bissau	Mali	Saudi Arabia
Singapore	Slovak Republic	Slovenia	Solomon Islands
Somalia	South Africa	Spain	Sri Lanka
St. Kitts and Nevis	St. Lucia	St. Vincent and the Grenadines	Sudan
Suriname	Sweden	Switzerland	Syria
Tajikistan	Tanzania	Thailand	Togo
Tonga	Trinidad and Tobago	Tunisia	Turkey
Turkmenistan	Uganda	Ukraine	United Arab Emirates
United Kingdom	United States	Uruguay	Uzbekistan
Vanuatu	Venezuela	Vietnam	Yemen
Zambia	Zimbabwe		

Figure A1: Number of Banking Crises

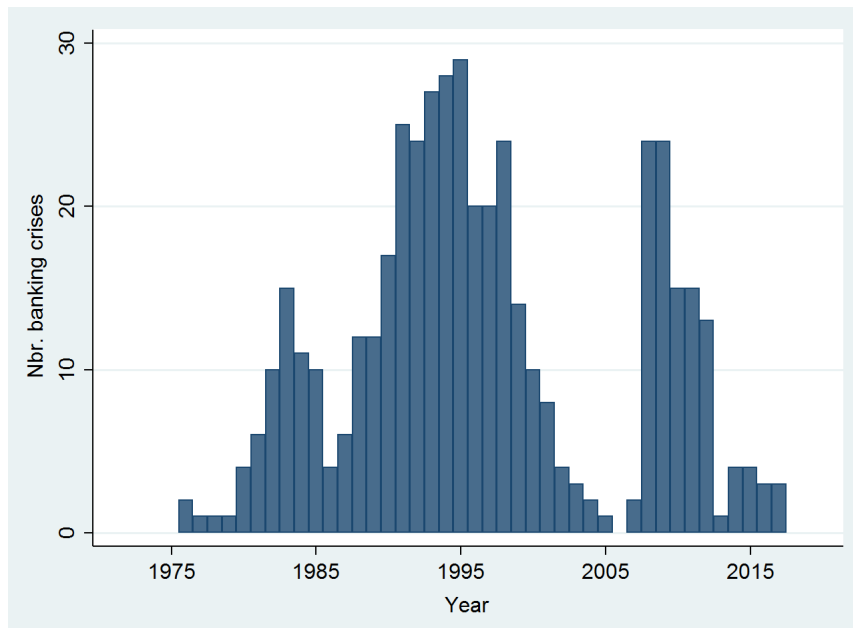


Table A2: Summary statistics of different national income shares

Variables	Obs	Mean	St. Dev	Min	Max
1- Top 1%	2,636	15.38	5.54	5.96	63.76
2- Top 10%	2,624	44.38	9.29	24.58	79.87
3- Top 20%	2,591	58.93	8.67	38.27	85.52
4- Middle 21%-79%	2,591	6.66	1.15	1.89	15.17
5- Bottom 20%	2,590	4.18	1.26	0.20	12.66
6 -Bottom 10%	2,584	2.88	1.38	0.01	11.53

Table A3: Banking crisis and the middle class

	Upper Middle class (51 to 79 percentiles) OLS	Lower Middle class (21 to 50 percentiles) OLS
Banking crisis	0.280* (0.051)	0.097** (0.043)
Observations	2566	2566
Countries in sample	140	140

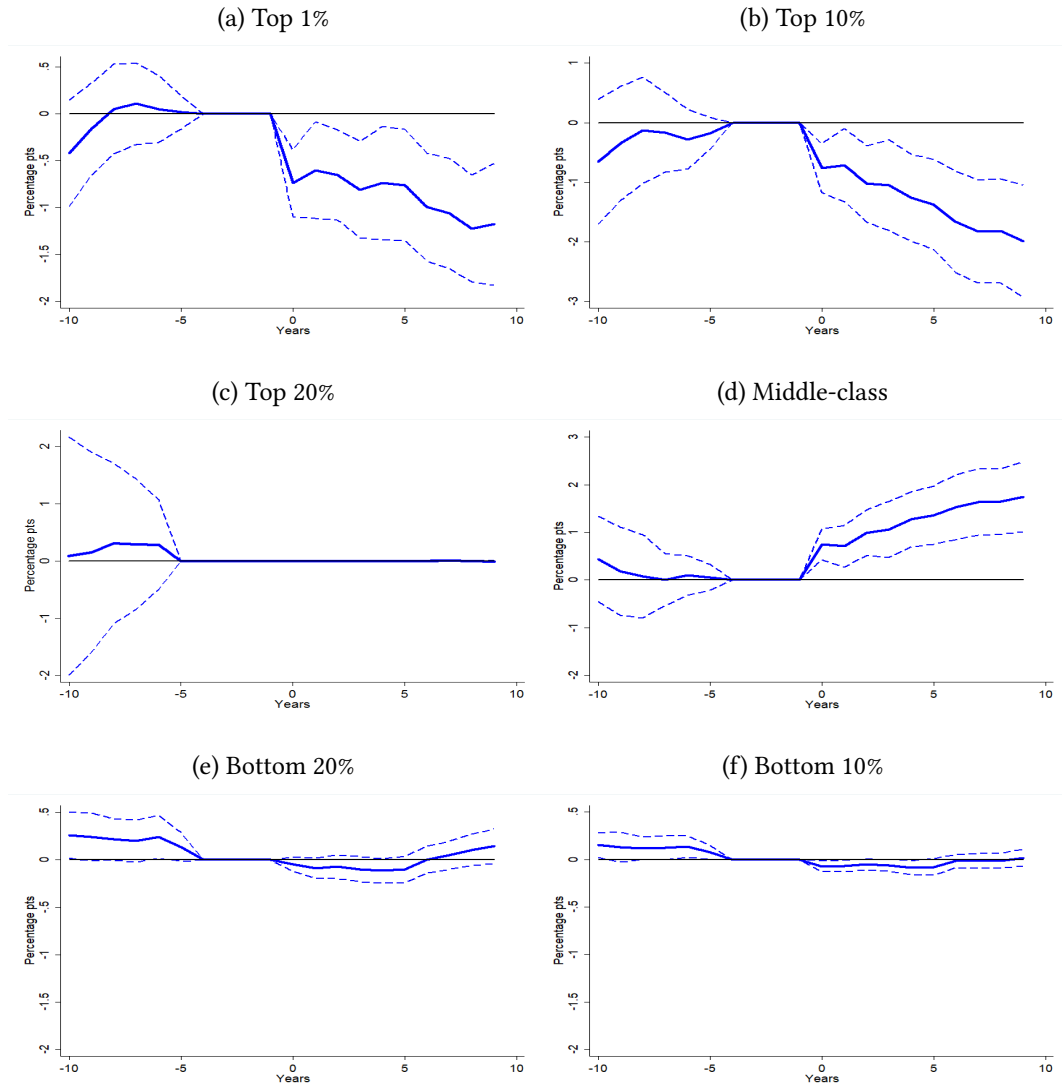
Note: This table shows the OLS results from the estimation of equation 1 for the national income shares held by the upper middle class, households positioned between P51 and P79 of the income ladder (column (1)) and by the lower middle class, households positioned between P20 and P50 of the income ladder (column (2)). A vector of control variables as well as country and time fixed effects are included, but not reported. Cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table A4: Banking crisis and the distribution of national income - Income percentile share ratio

	P1/P10	P1/P20	P1/P50	P1/P80	P20/P50	P20/P80	B20/B50	B10/B50	B10/B20
Banking Crisis - OLS	-0.005*** (0.002)	-0.004*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.004*** (0.001)	-0.002*** (0.001)	-0.015** (0.007)
Observations	2615	2566	2566	2566	2566	2566	2566	2566	2566
Countries in sample	140	140	140	140	140	140	140	140	140
Banking Crisis - IV	-0.017*** (0.006)	-0.014*** (0.005)	-0.010** (0.004)	-0.008** (0.003)	-0.009*** (0.002)	-0.010*** (0.003)	0.007 (0.007)	0.006 (0.005)	-0.000 (0.009)
Observations	2536	2487	2487	2487	2487	2487	2487	2487	2487
Countries in sample	137	137	137	137	137	137	137	137	137

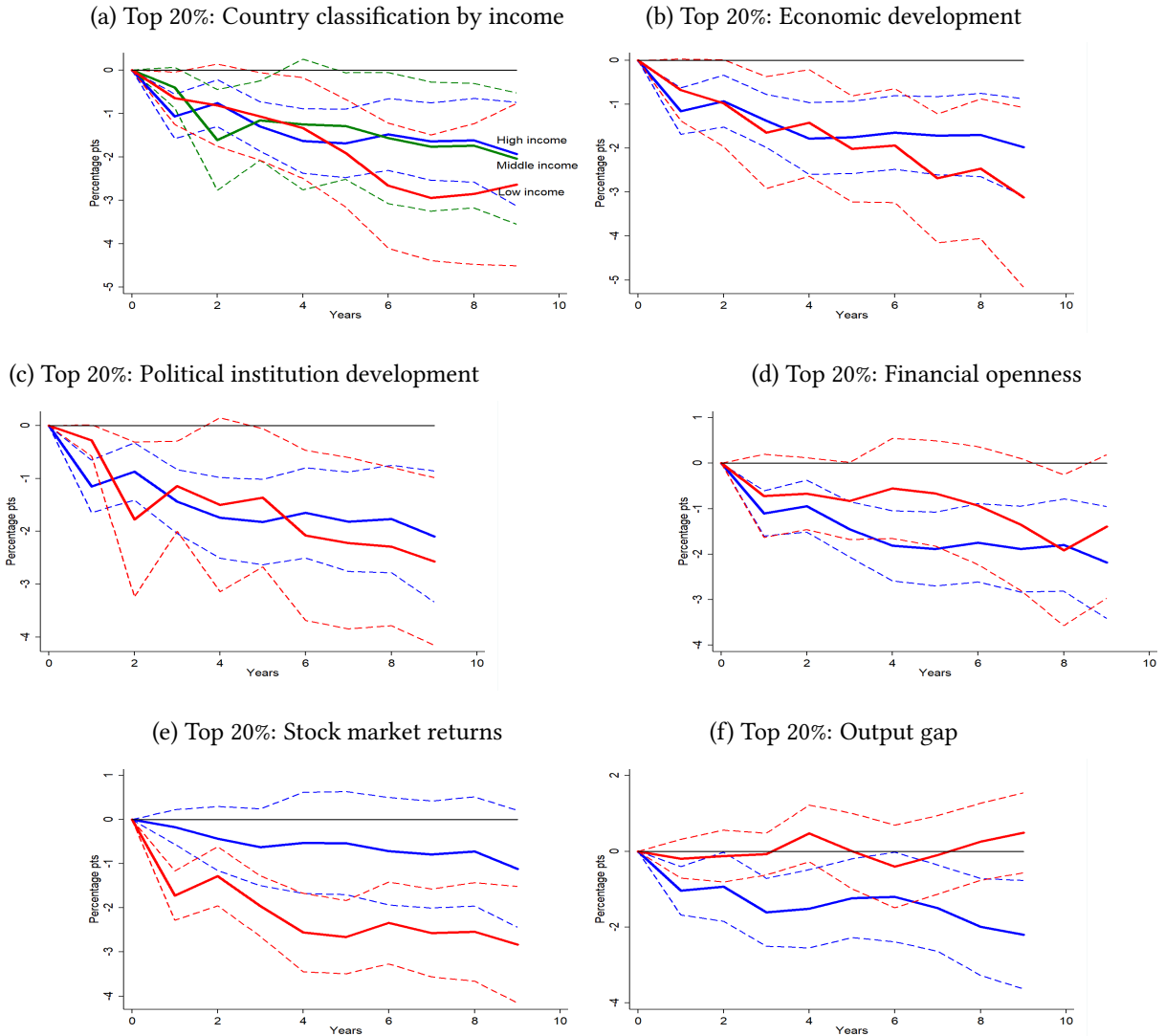
Note: This table reports the estimated effect of a banking crisis on different income percentile share ratios. We report estimates both from our OLS model (equation 2) and our IV model (equation 4). Cluster-robust standard errors are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Control variables and fixed effects are not reported.

Figure A2: ATET based on the regression adjustment estimator



Note: The figures show the over-time average treatment effects on the treated (ATET) of banking crises on the income share of different income groups. The ATET are obtained from the regression adjustment approach and are depicted by solid lines with 90% confidence bands based on cluster-robust standard error estimates in dashed lines.

Figure A3: Conditional ATET based on the doubly robust estimator



Note: The figures show the over-time conditional average treatment effects on the treated (ATET) of banking crises on the income share of P20. The ATET are obtained from the doubly robust estimator and are depicted by solid lines with 90% confidence bands based on cluster-robust standard error estimates in dashed lines. The red lines correspond to the estimates for a low level (regime with observations < empirical median) of the variable of interest, while the blue lines correspond to the estimates for a high level (regime with observations > empirical median)