ONLINE APPENDIX 3: ESTIMATION OF INCOME INEQUALITY OF GROUP 1

Supplementary material of: Astorga, Pablo. 2024. Revealing the diversity and complexity of long-term income inequality in Latin America: 1920-2011. *Journal of Economic History*, 84(4).

This appendix describes the procedure used to estimate Group 1's income inequality. It starts with an account of top income shares calculated with tax data. These shares are the basis for the calculation of beta coefficients and Fiscal Ginis for the top 1%, top 5% and top 10% of the distribution. The final step is the use the Fiscal Ginis to calculate income inequality in Group 1. The appendix ends with a sensitivity analysis.[1]

1. Top income shares using tax data

- *Argentina:* top1% shares (t1%), t0.5%, t0.1%, and t0.01% are available from 1932 to 2004, with interpolations in 1955, 1957, 1960, 1962-72, and 1974-96 in (Alvaredo 2010, Table 6.5). I assume that t1% between 1920 to 1931 are equal to the share in 1932, and between 2005 and 2011 are equal to the share in 2004.

  Income shares for the top5% are available in 1953, 54, 56, 59, 61, and 1997 from the same source. During 1932-1953, 1973, and 1997-2004 I grow t5% in line with a proxy series obtained by applying the t0.5%/t0.1% ratio to the t1% series, assuming that proportionality holds. In 1955, 1957, 1960, 1962-72, and 1974-96, t5% are interpolated.

  The top10% shares between 2000 and 2011 are sourced from the World Inequality Database (WID) - pre-tax national income concept. Shares between 1970 to 1996 are assumed equal to the 2000 share.

- *Brazil:* income shares are available for t1% and t0.1% in 1926-2011 from Souza (2018), with interpolations in 1929-32, 1961-62, 1988-95, 1999, 2001, 2003-05. And for t5% and t10% in 1969-2011, with interpolations in 1988-95, 1999, 2001, 2003-05.

  I assume that t5% and t10% between 1926 and 1968 grow in line with t1%; and that values between 1920 and 1925 are equal to the 1926 value.

- *Chile:* income shares are available for t10%, t5%, t1%, t0.5% t0.1%, t0.01% during 1964-2011 from Flores, Sanhuesa, Atria (2019), with interpolations in 1972, 75, 77, and 1982-90. Shares for t%5 are also interpolated in 1973-80.

- *Colombia:* income shares are available for t1%, t0.5%, t0.1%. t0.01% in 1993-2010 from Alvaredo & Londoño (2013, Table A4). I estimate a proxy series t5% by applying the

[1] I am most grateful to Facundo Alvaredo, Ignacio Flores, Marc Morgan, Javier Rodríguez Weber, and Pedro Souza for generous help with ideas and data.

t0.5%/t0.1% ratio to the t1% series, assuming that proportionality holds. Top 10% shares between 1993 and 2011 are sourced from WID - pre-tax national income.

- *Mexico*: t1%, t5%, t10% in 2012 are sourced from Bourguignon (2018, Table 1). Bourguignon includes two estimating options: the “Pareto adjusted” with a beta coefficient equal to 2.93; and the “mixed method” which combined adjusting the population above the 90th percentile in the original distribution with the “Pareto adjustment”. For my calculations I use the beta coefficients of the second option: 2.08 for the 99th percentile, 2.04 for the 95th percentile, and 2.67 for the 90th percentile.
- *Venezuela:* there are no estimates of the top income shares based on fiscal data. And any top shares estimates based on household budget surveys are grossly underestimated because they do not include property income (Maldonado 2021). Given these limitations, and for the sake of methodological consistency, I assume that t1%, t5%, and t10% around 2000 equal the simple averages of the corresponding income shares available for Argentina, Brazil, Chile and Colombia and Mexico.

2. Beta coefficients

Following Alvaredo (2011), the beta coefficients are calculated from pairs of top shares as:

(1) $\beta=\alpha/(\alpha-1)$, with Pareto coefficient $\alpha = 1/[1-\log(Sx\%/Sy\%)/\log(x/y)]$.

For example, for the pair t1% and t0.1%, Sx% is the top1% share, Sy% the top0.1% share, and log(x/y) is log(1/0.1). Different pairs usually result in different betas.[2] For each country, I calculate simple average betas (*bx%*) for t10%, t5%, and t1% as follows:

- *Argentina: b10%* uses a single pair t10%&t1%;[3] *b5%* is the average of pairs t5%&t1% and t5%&t0.1%; *b1%* in the average of pairs t1%&t0.5%, t1%&t0.1%, and t1%&t0.01%. The resulting beta ratios b10%/b5% and b5%/b1% are 0.75 (coefficient of variation = 6.3%) and 1.04 (= 4.4%) respectively.
- *Brazil: b10%* is the average t10%&t5% and t10%&t1%; *b5%* is the average of t5%&t1% and t5%&t0.1%; *b1%* is a single pair t1%&t0.1%. The resulting beta ratios b10%/b5% and b5%/b1% are 1.02 (coefficient of variation = 5.3%) and 1.03 (= 9.6%) respectively.
- *Chile: b10%* is a single pair t10%&t1%; *b5%* is the average of pairs t5%&t1%, t5%&t0.5%, and t5%&t0.1%; *b1%* is the average t1%&t0.1% and t1%&t0.01%. The resulting beta ratios b10%/b5% and b5%/b1% are 0.94 (coefficient of variation = 11.7%) and 1.31 (= 13.4%) respectively.

---

2 Similar betas with ratios close to unity indicates that the empirical distribution follows a Pareto form which is characterised by a constant beta.

3 Because implausible betas for the pair t10%&t5% are excluded in the calculations.

- *Colombia: b10%* is a single pair t10%&t1% (both income shares from WID); *b5%* is the average of pairs t5%&t1%, t5%&t0.5% and t5%&t0.1%; *b1%* is the average of pairs t1%&t0.1% and t1%&t0.01%. The resulting beta ratios b10%/b5% and b5%/b1% are 1.05 (coefficient of variation = 9.1%) and 0.96 (= 4.7%) respectively.
- *Mexico:* estimates in 2012 from Bourguignon (2018). I use his mixed-method betas (see above): *b10%*=2.64; *b5%*=2.04; *b1%*=2.08. The ratios b10%/b5% and b5%/b1% are 1.29 and 0.98 respectively.
- *Venezuela:* there are no estimates of the top income shares based on fiscal data (see above). Facing this limitation, I assume *b10%*, *b5%*, and *b1%* around 2000 as the simple average of those betas calculated for Argentina, Brazil, Chile, Colombia and Mexico.

3. Fiscal Ginis

Following Alvaredo (2011), under the assumption that the income distribution at the top follows the Pareto form, the corresponding Fiscal Ginis (FGini*tx%*) for t1%, t5% and t10% are calculated as:

(2) FGini*tx%* = (*bx*%-1)/(*bx*%+1).

Additionally, FGini*t3%* and FGini*t7.5%* are obtained as a simple average of the adjacent Fiscal Ginis. For example, FGini*t3%* = (FGini*t1%* + FGini*t5%)*/2.

For Chile and Colombia, I rely on proxy Ginis measuring income inequality among high earners to extend backwards the Fiscal Ginis. In Chile I use changes in a Gini series of employers from Rodríguez-Weber (2014) between 1929 and 1963. Values from 1920 to 1928 are assumed equal to those in 1929. In Colombia I use changes in Gini series of employers and landlords (weighted by their employment shares) between 1938-1988 from Londoño (1995). To splice this series with my Fiscal Ginis, I assume that Londoño's Ginis in 1988-1993 are equal to the 1988 value. Fiscal Ginis from 1920 to 1937 are assumed equal to the value in 1938, and those in 2011 to the value in 2010.

4. Group 1's income Ginis

The final step is to calculate income Ginis for Group 1 (Gg_1)[4] by matching in each year between 1920 and 2011 the EAP shares of the top group (e_1) with the appropriate Fiscal Gini out of the five options ranging from top1% to top10% as follows:

- those EAP shares in the interval 0%< e_1 <2% are assigned values of FGini*t1%*;
- those in 2% ≤ e_1 < 4% are matched with FGini*t3%*;
- those in 4% ≤ e_1 < 6.5% with FGini*t5%*;
- those in 6.5% ≤ e_1 < 8.5% with FGini*t7.5%*;
- and those in 8.5% < e_1 ≤ 12% or higher with FGini*t10%*.

4 See equation 3 in the Appendix of the paper.

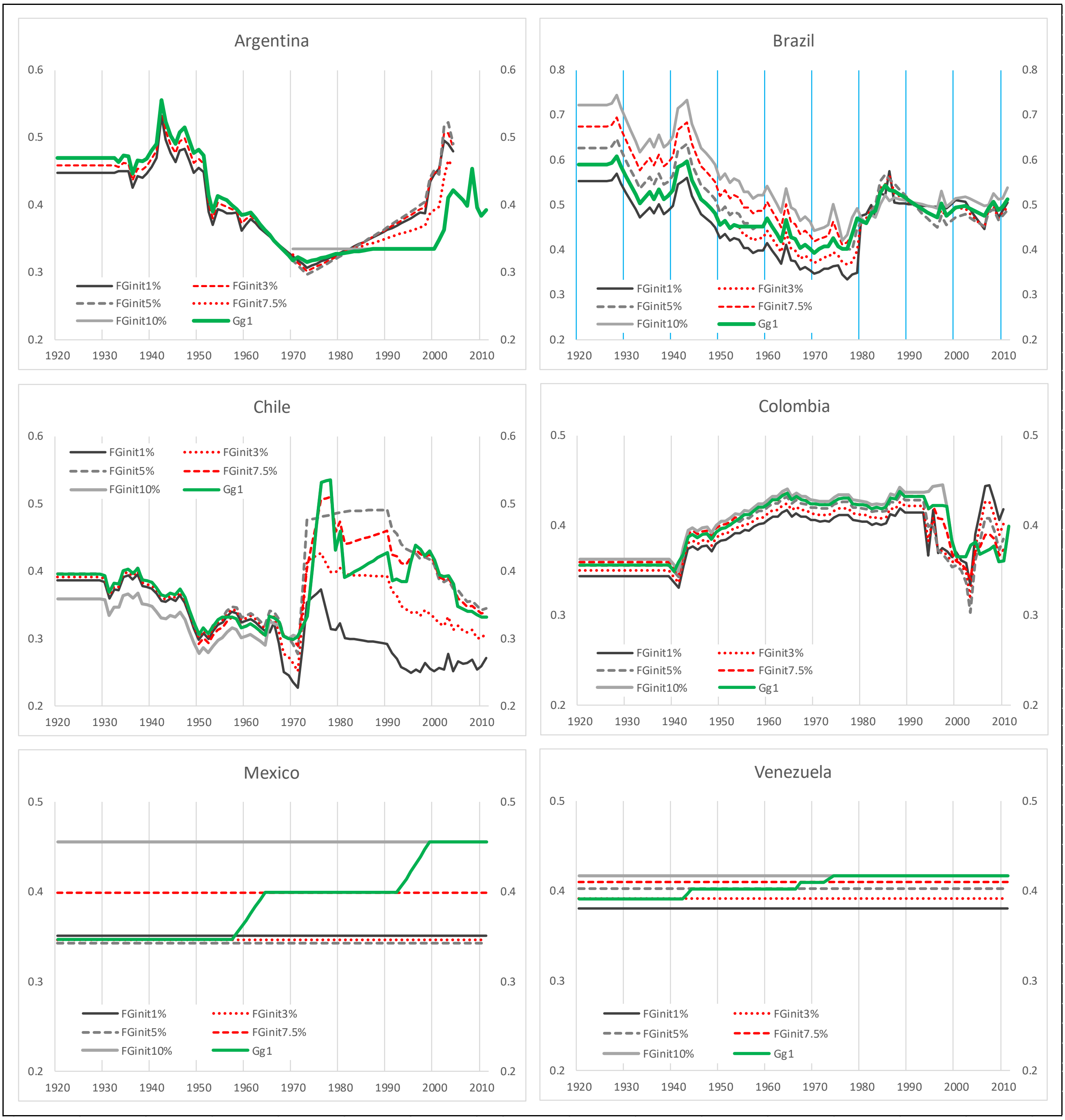
Figure OA3.1 shows the resulting Gg_1 by country. In practice, the lowest e_1 in the six countries is 2.9% in Mexico c.1940 and the highest 14.5% in Chile c.2005. I use interpolations to smooth out the transitions between adjacent Fiscal Ginis in Mexico during 1957-64 and 1992-99; and in Venezuela in 1943 and 1973.

5. Sensitivity analysis

There are two estimation issues when calculating Gg_1. First, there are only series of income dispersion at the top fractiles of the distribution for Argentina, Brazil, Chile, and Colombia with differences is historical coverage and, depending on the country, with significant data gaps. Secondly, estimates of Group 1's income inequality are based on Fiscal Ginis for the top1%, top5% and top10% but, in years where such Ginis are not available, I am assuming that they remain constant. Then, a pertinent question is: What impact would have any misrepresentation of the "true income inequality" of Group 1 on the *G4* and *G4W* series? To answer this, I perform two robustness checks.

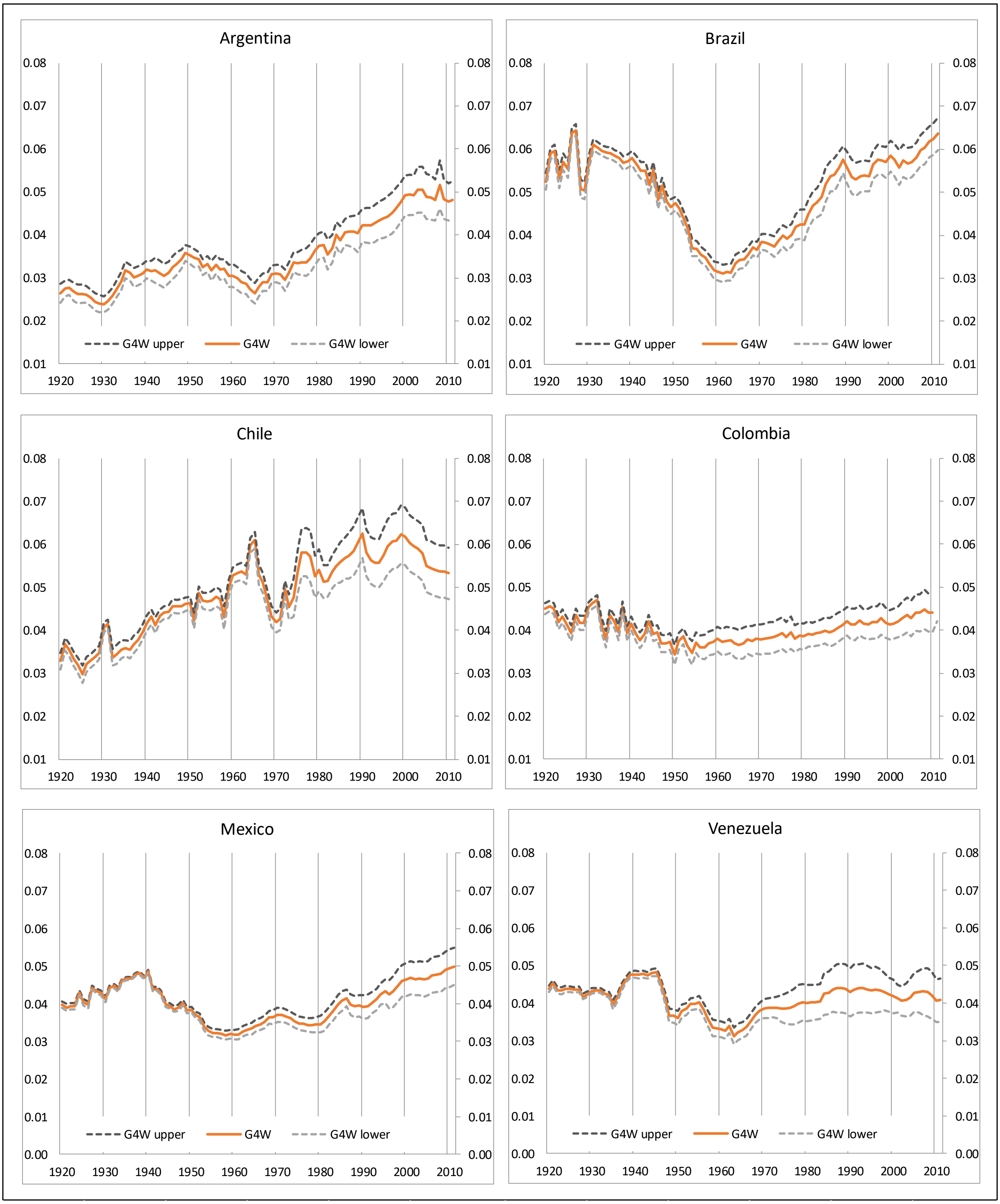
- The first one tests the impact on the two occupational Ginis of using ±20% upper and lower bounds in Gg_1. The outcome of this exercise is shown in Figure AO3.2 (*GW4*) and Figure AO3.3 (*G4*). Although there are relatively significant deviations in the estimation of *G4W* (e.g., in Argentina and Chile post 1980), they have little impact on *G4*. This result is to be expected (Alvaredo 2011) because, although the estimated income inequality in the top group is the largest of the four groups, its contribution to *G4* is relatively low owing to its reduced EAP share.
- The second check focuses on Mexico and Venezuela. In both cases I assume that inequality remains equal to values calculated c.2012 and c.2000 respectively, which could have serious accuracy implications. I perform a sensitivity exercise using beta=3 (high) and beta=2 (low) in the calculation of Gg_1. Figure OA3.4 shows that although there are significant deviations for *G4W* (e.g., post 1980 in Mexico and post 1970 in Venezuela), the impact on *G4* is very limited.

FIGURE OA3.1: FISCAL GINIS AND GROUP1'S INCOME GINIS



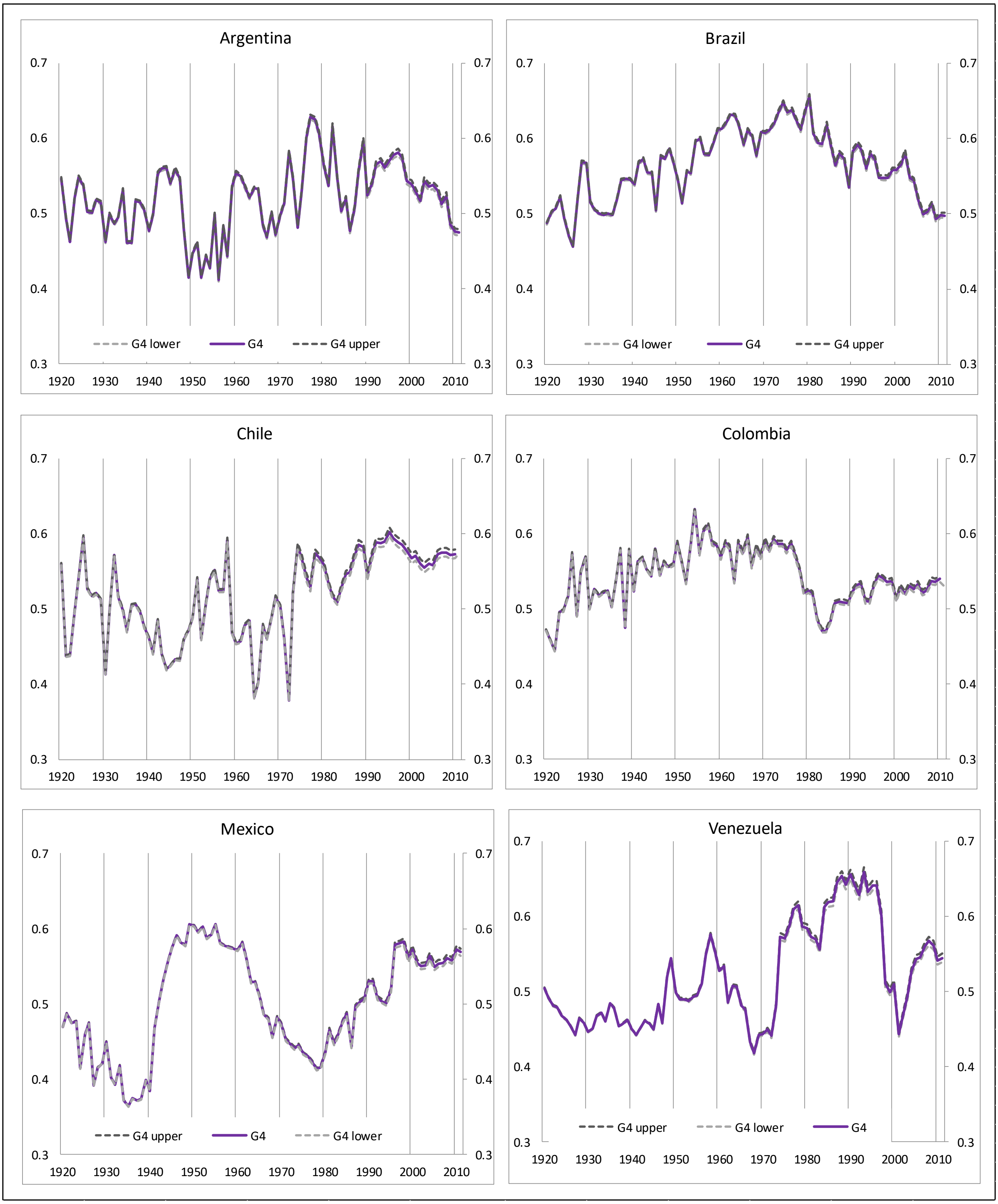
Notes: FGinit$x$% = Fiscal Gini of the top$x$%; Gg_1 = Group 1's income Gini.

FIGURE OA3.2: SENSITIVITY ANALYSIS ON G4W (WITH UPPER=+20% & LOWER=-20% LIMITS)



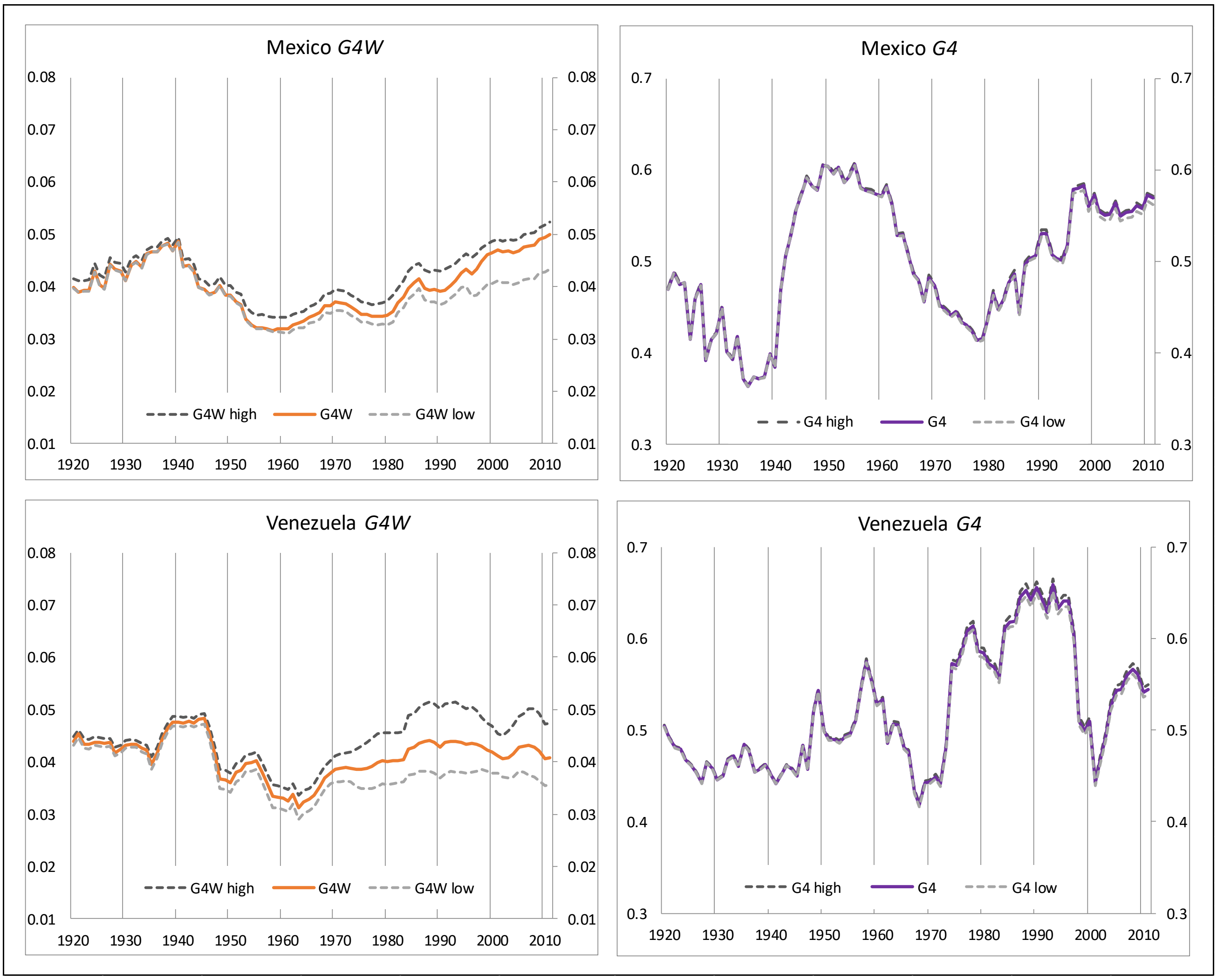
Notes: *G4W* = within-groups income Gini including all four groups; *G4W upper* = the Gg_1 component equals 20% above *G4W* (baseline); *G4W low* = the Gg_1 component equals 20% below *G4W*.

FIGURE OA3.3: SENSITIVITY ANALYSIS ON G4 (WITH UPPER=+20% & LOWER=-20% LIMITS)



Notes: *G4* = Overall Gini; *G4 upper* = calculated with *G4W upper* (see Figure OA3.2); *G4 lower* = calculated with *G4W lower*.

FIGURE OA3.4: SENSITIVITY ANALYSIS ON G4W & G4 WITH ALTERNATIVE BETAS (HIGH=3; LOW=2) IN MEXICO AND VENEZUELA



Notes: *G4W* = within-groups income Gini including all four groups; *G4W high* = the Gg_1 component is calculated with beta=3; *G4W low* = the Gg_1 component is calculated with beta=2; *G4* = Overall Gini; *G4 high* = calculated with *G4W high*; *G4 low* = calculated with *G4W low*.

References

Alvaredo F. 2011. A note on the relationship between top income shares and the Gini coefficient. *Economics Letters*, 110(3), 274-277.

Alvaredo F. 2010. The rich in Argentina over the twentieth century 1932-2004. In Atkinson, A. and Piketty, T. (eds.) *Top Incomes vol. II: A Global Perspective*. Oxford University Press.

Alvaredo F, Londoño Vélez J. 2013. High income and personal taxation in a developing economy: Colombia 1993-2010. Commitment to Equity working paper 12, May.

Bourguignon F. 2018. Simple adjustments of observed distributions for missing income and missing people. *J Econ Inequal*, 16, 171-188.

Flores I, Sanhueza C, Atria J, Mayer R. 2019. Top incomes in Chile: A historical perspective on Income Inequality, 1964-2017. *Review of Income and Wealth*, 66(4), 850-874.

Londoño JL. 1995. *Distribución del Ingreso y Desarrollo Económico: Colombia en el Siglo XX*. Santafé de Bogotá: Tercer Mundo Editores.

Maldonado LJ. 2021. Decomposing the gender pay gap in the formal sector in Venezuela: a microdata analysis 1985–2015. *Applied Economics Letters*, 28(14), 1145-1151.

Rodríguez Weber J. 2014, *La Economía Política de la Desigualdad de Ingreso en Chile*, Doctoral dissertation, Universidad de la República de Uruguay.

Souza PHGF. 2018. A history of inequality: top incomes in Brazil, 1926-2015. *Research in Social Stratification and Mobility*, 57, 35–45.

World Inequality Database (WID). Last accessed on July 2022 from: http://wid.world/