Income Shock and Redistribution Under Misreporting

Gaspar Alarcón
Matías Ustares

Quantitative Macroeconomic Theory
Washington University in St Louis

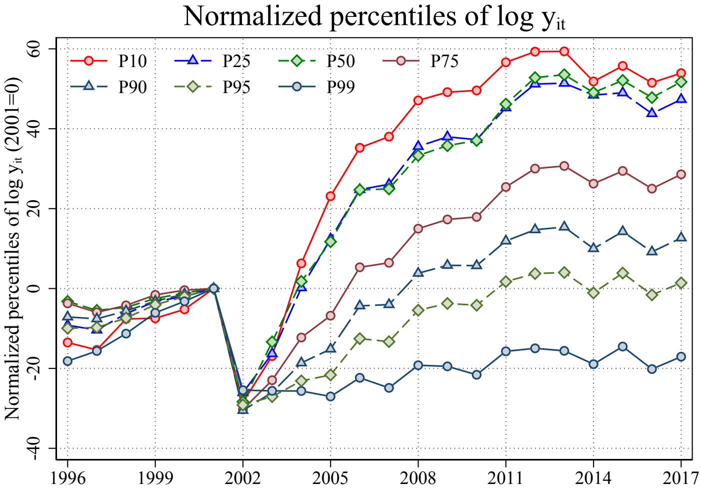
December 6, 2021

Outline

Outline

Argentina

25% drop during 2002 devaluation and economic crisis, and rapid bottom-driven growth after 2002 devaluation. Details



Questions

- In the 2001/2002 crisis, wages fell significantly, in real terms, for most workers in the formal sector.
- After the crisis, we observe that the recovery presents different heterogeneities. The lowest percentiles recover the pre-crisis level in just three years, while the highest do so in up to ten years, or the very top, never recover their previous level.
- Could misreported wages reflect an illusory distribution improvement?

Outline

Data source & Sample Selection

Sistema Integrado Previsional Argentino (SIPA)

- Matched employer-employee monthly panel data (1996-2017)
- Covers all *formal* workers (private and public)
- Data on 4.5 million workers per year

Sample Selection

- Earnings include salary, bonus, vacation, 13th salary
- Compute total annual earnings, y_{it}, for person i in year t
- Drop observations with incomes below threshold

$$y_{it} \leq \frac{1}{2} \times 3 \text{ months } \times \text{ National } MW_t$$

Outline

Model

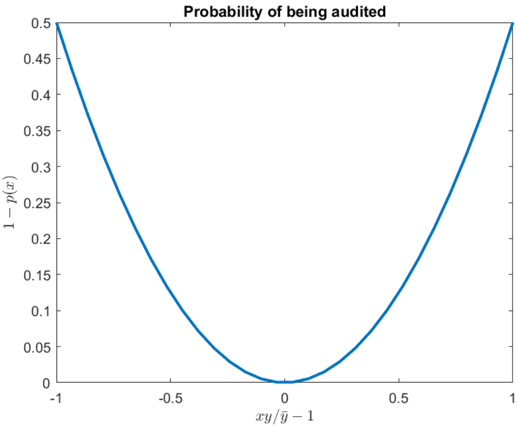
$$V(\varepsilon, z, \bar{y}) = \max_{x} \; p(x)\left\{u\left[y - T(x) - F(x)\right] + \beta E_{\varepsilon}\left[V(\varepsilon', z', xy)\right]\right\}$$
$$+ \left[1 - p(x)\right]\left\{u\left[(1-\delta)y - F(x)\right] + \beta E_{\varepsilon}\left[V(\varepsilon', z', y)\right]\right\}$$

subject to

$$\begin{aligned} &x \in [0,1], \\ &T(x) = \tau_{\Delta}(xy - \bar{y}) + \tau_y xy, \\ &0 \leq y - F(x) - \max\{\delta y, T(x)\}, \end{aligned}$$

where $y = e^{\varepsilon + z}$, $\varepsilon' - \rho\varepsilon \sim N(0, \sigma_{\varepsilon}^2)$, and $z' \sim N(0, \sigma_z^2)$. Utility is $u(c) = c^{1-\gamma}/(1-\gamma)$.

$$1 - p(x) = \min\left\{1, \frac{1}{2}\left(\frac{xy - \bar{y}}{\bar{y}}\right)^2\right\}$$



Algorithm

Discretize ε (using Tauchen) and z.

1. Create a grid for x: $\{x_1, x_2, \ldots, x_N\}$ with $x_1 = 0$ and $x_N = 1$.
2. Guess V^0 in the state grid $(\varepsilon, z, \bar{y})$.
3. For every $(\varepsilon, z, \bar{y})$, compute

$$\begin{aligned} M_i = {} & p(x_i)\left\{u\left[y - T(x_i) - F(x_i)\right] + \beta E_\varepsilon\left[V^0(\varepsilon', z', x_i y)\right]\right\} \\ & + \left[1 - p(x_i)\right]\left\{u\left[(1-\delta)y - F(x_i)\right] + \beta E_\varepsilon\left[V^0(\varepsilon', z', y)\right]\right\} \end{aligned}$$

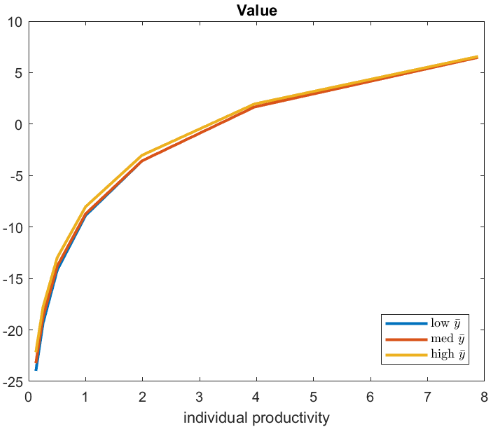
   for every x_i in the grid. To compute $E_\varepsilon V^0(\varepsilon', z', x_i y)$, we interpolate along state $\bar{y}$.
4. Set $V^1(\varepsilon, z, \bar{y}) = \max\{M_i\}_{i=1}^N$.
5. If V^1 is close to V^0, stop. Otherwise, set $V^0 = V^1$ and go back to step 3.

Calibration

Discount factor	β	0.96
Elasticity of IS	γ	2.00
Persistance of ε	ρ	0.90
Std. Dev. $\varepsilon' - \rho\varepsilon$	σ_ε	0.30
Std. Dev. z	σ_z	0.04
Tax/subsidy rate on $(xy - \bar{y})$	τ_Δ	0.20
Tax rate on xy	τ_y	0.30
Misreporting fine rate	δ	0.50

Checking the answer

Fix $F(x) = 20 > 0$ for $x \neq 1$ and $F(1) = 0$.

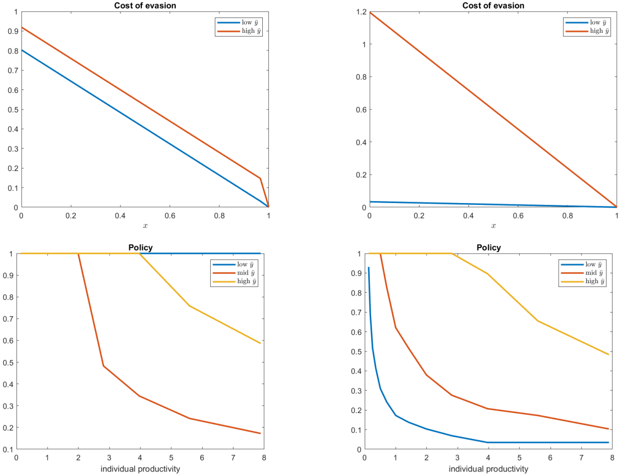


The policy is $x(\varepsilon, z, \bar{y}) = 1$ for all $(\varepsilon, z, \bar{y})$.

Comparing $F(x)$

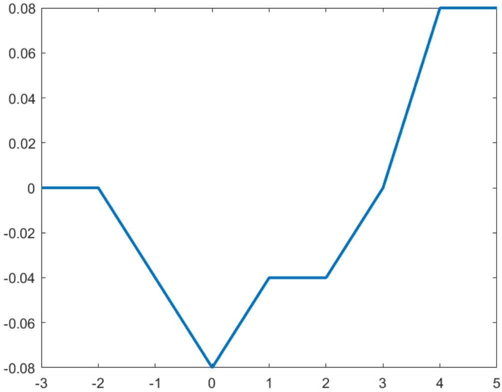
$$F_1 = 0.8(1-x) + 0.02y \qquad F_2 = 0.2(1-x)y$$

Value Functions F_3



Outline

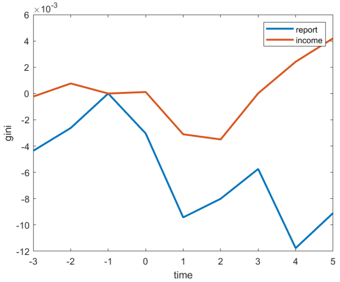
Aggregate shock z



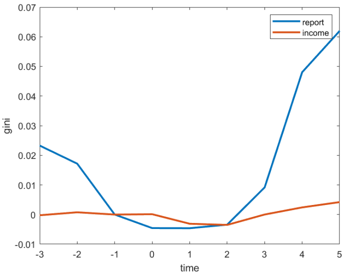
Gini coefficient

Simulation with 10,000 individuals and 500 periods.

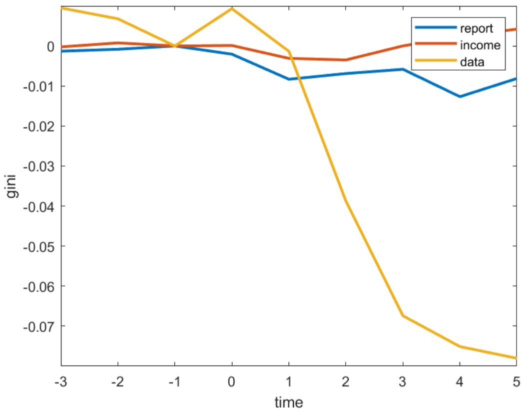
F_1



F_2



Gini under F_1 and data



Statistics under cost F_1

The following table presents the variation of the main statistics between the period before the crisis (1998-2000 in the data and [-3, -1] in the model]) and after (2001-2005 and [0, 3]). Details

Stats	Δ Data	Δ Reported y_{it} (Model)	Δ Effective y_{it} (Model)
Gini	-3.79%	-1.08%	-0.27%
CV	-3.11%	-1.55%	-0.18%
Mean	-0.87%	-3.26%	-2.76%
Median	-0.61%	-3.92%	-3.92%

Outline

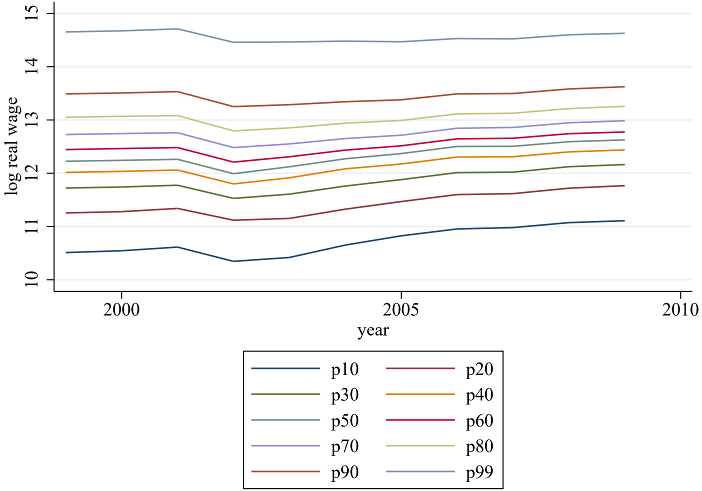
Conclusions

- In our analysis, the reported income shows less inequality than the effective income distribution after the shock.
- The distribution of the reported income is more sensitive to the aggregate shock than effective income.
- Misreporting magnifies changes in the income distribution.

THANK YOU!

Percentiles of log y_{it}

We can observe the heterogeneity in the recovery by plotting the percentiles without normalizing them. Return



Statistics using reports under F_1

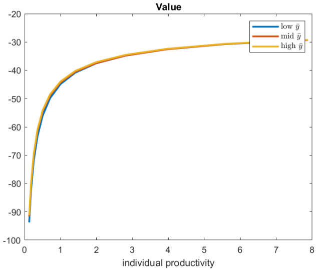
The following table presents the absolute values of the main stats between the period before the crisis (1998-2000 in the data and [-3, -1] in the model]) and after (2001-2005 and [0, 3]). Return

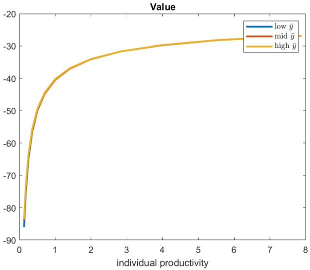
	Pre-shock [-3, -1]		Pos-shock [0, - 3]	
Stat	Data ('98 - '00)	Model	Data ('01 - '04)	Model
Gini	0.52	0.36	0.50	0.36
CV	0.93	0.71	0.90	0.71
Mean	12.14	1.20	12.03	1.17
Median	12.22	1	12.15	0.96

Value functions

$$F_1 = 0.8(1-x) + 0.02y \qquad F_2 = 0.2(1-x)y$$

Return





Cost $F_3 = \frac{0.2}{x} + 0.01y$

Return