

Using Microsimulations to Estimate Historical Income Distributions at High

Frequency:

Italy, 1861-2021

Brian A'Hearn (University of Oxford)

Nicola Amendola (University of Rome Tor Vergata)

Federico Belotti (University of Rome Tor Vergata)

Giulia Mancini (University of Sassari)

Giovanni Vecchi (University of Rome Tor Vergata)

During the past two decades or so, an unprecedented amount of effort has been dedicated to reconstructing the dynamics of inequality and living standards in the long run, through a variety of approaches and data sources. The work has clearly paid off, in the form of a wealth of empirical evidence and historiographical insight (Rossi, Toniolo and Vecchi, 2001; Saez and Piketty, 2003; Prados de la Escosura, 2008; Atkinson, Piketty and Saez, 2011; Lindert and Williamson, 2016; Scheidel, 2018; Chancel and Piketty, 2021; Alfani, 2023). This new knowledge has been assembled by resourcefully working around data limitations and making some necessary compromises: strong assumptions and imputations compensate for a lack of quality household-level information; second-best indicators are used when estimating a full array of inequality indices is impossible; and trade-offs are made between accuracy and frequency of estimates. As a result, plenty of challenges still lie ahead.

This paper pushes against some of these challenges by focusing on Italy, where first-generation estimates of the income distribution at ten-year intervals, built on roughly 25,000 household budgets from 1861-1961, have laid the groundwork for the study of long-run poverty and inequality (Vecchi 2011, 2017). On top of this, a range of proxy indicators of inequality have been constructed (Sylos Labini 1974; Zamagni 1980; Gabbuti, 2020a, 2020b; Gómez León and Gabbuti 2022). Our contribution is

to present second-generation estimates of the full household income distribution in Italy, at annual frequency, from 1861 to the present. This is an ambitious task, one that goes well beyond the estimation of summary measures of inequality and poverty: generating such high-resolution estimates requires new data and new methods to fill in the gaps.

On the data front, this work relies on an expanded version of the Historical Household Budgets (HHB) dataset, a collection of harmonized household budgets extracted from heterogeneous historical sources (A'Hearn, Amendola, and Vecchi, 2016). The dataset underlying Vecchi (2011, 2017) has been augmented, now including additional budgets from aristocratic families in the upper income brackets (with the understanding that nobility does not necessarily equate to wealth) and from applicants for government assistance in the lower brackets, for example. More importantly, new types of data have become available, including some derived from administrative sources related to the *tassa del focatico* (a household tax implemented on a municipal basis starting in the late 1860s) and others from company archives, which often preserve personnel files rich in information about workers' households. An important role is also played by sources containing tabulated aggregates, or “grouped data,” which bear some similarity to the social tables in the style of Milanovic, Lindert, and Williamson (2011). As a result, the second-generation HHB database currently comprises about 350,000 household-level observations for the period 1861-1971, more than ten-fold its previous size.

As for methods, the paper taps from a literature that has developed separately from economic history, but can be usefully adapted to the study of the past: microsimulations. Traditionally used to “look forward” with the primary aim of informing policymaking, these tools can be applied to “look backward”, supporting economic historical analysis. Since Orcutt (1957) introduced the idea that predictions about macroeconomic aspects – such as changes in GDP, unemployment rates, and other economic indicators – are more accurately obtained by aggregating information on the behavior of micro units (individuals, households, firms, labor unions, etc.) rather than relying solely on aggregate variables, microsimulation methods have gained traction. Their use expanded significantly in the 1980s and 1990s alongside advances in computing power. The classic application of these

methods is in welfare and redistributive design, with tax-benefit simulation models being a primary example (Sutherland 1995; Colombino 2015). International organizations such as the World Bank and the International Monetary Fund have contributed significantly to the literature on forecasting poverty and inequality measures (Ravallion 2013; Bourguignon and Bussolo 2013; Yoshida et al. 2014; Lackner et al. 2022). In recent years, several events have significantly increased the demand for microsimulations: inflationary surges, the COVID-19 pandemic, wars, natural disasters, accelerating migration flows, and much more. These events have heightened the need to understand the impact of crisis-induced changes on income distribution and their broader economic effects on society.

The current version of the paper focuses on so-called “static models”, which simulate shifts in household incomes, but do not account for agents’ behavioral responses (*e.g.* movements in and out of the labor force, or across economic sectors; Li et al. 2014). These simulations are used to fill in the gaps (ranging from one to three years) between benchmark years selected from the historical household budget dataset, when income distributions can be reconstructed based on genuine data.

Although the reconstructions presented in this paper rely on a database that is still under construction, and the results are highly preliminary, the relevance of the conceptual and methodological framework extends beyond the specific empirical case under examination. Arguably, what is feasible for Italy is achievable for a number of other countries traditionally richer in historical data; microsimulations, even in their simplest forms, are a valuable tool that allows economic historians to explore the past using a modern, tested and often reliable machinery.

References

A'Hearn, B., Amendola, N., & Vecchi, G. (2016). On historical household budgets. *Rivista di storia economica*, 32(2), 137-176.

Alfani, G. (2023). *As Gods Among Men: A History of the Rich in the West*. Princeton University Press.

Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top incomes in the long run of history. *Journal of economic literature*, 49(1), 3-71.

Bourguignon, F., & Bussolo, M. (2013). Income distribution in computable general equilibrium modeling. In P. B. Dixon and D. Jorgenson (eds.), *Handbook of computable general equilibrium modeling, Vol. 1*. Elsevier.

Chancel, L., and Piketty, T. (2021). Global income inequality, 1820–2020: the persistence and mutation of extreme inequality. *Journal of the European Economic Association*, 19(6), 3025-3062.

Colombino, U. 2015. "The ISTAT microsimulation models." *Rivista di statistica ufficiale* 2: 5-16.

Gabbuti, G. (2020a). 'A Noi! Income Inequality and Italian Fascism: Evidence from Factor and Top Income Shares', Oxford Economic and Social History Working Papers, 177.

Gabbuti, G. (2020b). 'Labor Shares and Inequality: Insights from Italian Economic History, 1895-1970', European Review of Economic History, doi:10.1093/ereh/heaa010.

Gómez León, M., and G. Gabbuti (2022). Wars, Depression, and Fascism: Income Inequality in Italy, 1900-1950. LEM Working Paper Series 2022/26.

Lakner, C., Mahler, D. G., Negre, M., and Prydz, E. B. (2022). How much does reducing inequality matter for global poverty?. *The Journal of Economic Inequality*, 20(3), 559-585.

Li, J., C. O'Donoghue, J. Loughrey, and A. Harding. 2014. "Static Models" in C. O'Donoghue (ed.) *Handbook of Microsimulation Modelling*, pp. 47-75, Emerald.

Lindert, P. H., and Williamson, J. G. (2016). *Unequal Gains: American Growth and Inequality since 1700*. Princeton University Press.

Milanovic, B., Lindert, P. H., and Williamson, J. G. (2011). Pre-industrial inequality. *The economic journal*, 121(551), 255-272.

- Orcutt, G. H. 1957. "A New Type of Socio-Economic System." *The Review of Economics and Statistics* 39, no. 2: 116–123.
- Piketty, T., and Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly journal of economics*, 118(1), 1-41.
- Prados de la Escosura, L. (2008). Inequality, poverty and the Kuznets curve in Spain, 1850–2000. *European Review of Economic History*, 12(3), 287-324.
- Ravallion, M. (2013). How long will it take to lift one billion people out of poverty?. *The World Bank Research Observer*, 28(2), 139-158.
- Rossi, N., G. Toniolo, and G. Vecchi. 2001. "Is the Kuznets curve still alive? Evidence from Italian household budgets, 1881–1961.?" *The Journal of Economic History* 61, no. 4: 904-925.
- Scheidel, W. (2018). *The great leveler: Violence and the history of inequality from the stone age to the twenty-first century*. Princeton University Press.
- Sutherland, H. 1995. "Static microsimulation models in Europe: A survey." Cambridge Working Papers in Economics 9523, University of Cambridge.
- Vecchi, G. (2011). *In ricchezza e in povertà: il benessere degli italiani dall'Unità a oggi*. Bologna : Il Mulino.
- Vecchi, G. (2017). *Measuring Wellbeing: A History of Italian Living Standards*. Oxford-New York: Oxford University Press.
- Yoshida, N., Uematsu, H. and Sobrado, C. E. (2014). Is Extreme Poverty Going to End? An Analytical Framework to Evaluate Progress in Ending Extreme Poverty World Bank Policy Research Working Paper No. 6740.
- Zamagni, V. (1980). *Distribuzione Del Reddito E Classi Sociali Nell'Italia Fra Le due Guerre*. *Annali Della Fondazione Feltrinelli*, vol. 20/(1980), pp. 17-50.