On Historical Household Budgets

Brian A’Hearn\textsuperscript{1}, Nicola Amendola\textsuperscript{2}, and Giovanni Vecchi\textsuperscript{3*}

Working Paper No. 45

May 2016

\textbf{Abstract}

The paper argues that household budgets are the best starting point for investigating a number of big questions related to the evolution of the living standards during the last two-three centuries. If one knows where to look, historical family budgets are more abundant than might be suspected. And statistical techniques have been developed to handle the associated problems of small, incomplete, and unrepresentative samples. We introduce the Historical Household Budgets (HHB) Project, aimed at gathering data and sources, but also at creating an informational infrastructure that provides i) reliable storage and easy access to historical family budget data, along with ii) tools to configure the data as it is entered so as to harmonise it with present-day surveys.

\textsuperscript{1} Pembroke College, University of Oxford
\textsuperscript{2} Dept. of Economics and Finance, University of Rome “Tor Vergata”
\textsuperscript{3} Corresponding author. Dept. of Economics and Finance, University of Rome “Tor Vergata”
giovanni.vecchi@uniroma2.it

\* The paper is part of a broader research project supported by the Institute for New Economic Thinking (INET), Grant #INO15-00026. Details are available at \url{http://hhbproject.com}. 
Keywords: household budgets; household budget surveys; living standards; inequality; poverty; survey; globalization; purchasing power parities; grouped data; post-stratification.

JEL classification: N30, I31, I32, C81, C83, D60, D63, O12, O15.
1 Introduction

Like a fairy-tale giant lifting the roof of a house to peer inside, economic historians wish they could see into the lives of families of the past. For important questions can only be answered with detailed evidence about individual households. This is obvious when interest centers on the intra-household allocation of resources, that is, questions of whose consumption was protected when times were tough: the primary breadwinner, the children, males? It is equally true of other questions naturally posed at the level of the household. Was poverty a feature of the life cycle, something episodic, or a chronic condition? How did parents choose between their children’s involvement in education, home production, and outside employment? Who benefited from rural-urban migration? On what margins did families react to shocks like a decline in wages: hours worked, home production, quality of diet or dwelling, wealth accumulation?

Convincing answers even to questions of a more macroeconomic nature require disaggregated welfare indicators. We are increasingly aware that the great forces of our era – globalisation and technical change – create losers as well as winners (O’Rourke and Williamson, 1999). We are increasingly concerned about rising inequality and its political implications (Bourguignon, 2015; Milanovic, 2016). We worry about future trends and whether policy is capable of making a difference (Piketty, 2014; Atkinson, 2015). This is true within and between countries, for the rich world and developing nations (Deaton, 2013).

Modern economists take household-level data for granted, but do they exist for times past? In this article we show that historical data on families are more than just a fairy tale. Household budgets, in particular, are available for more countries, for a longer span of time (back to the mid-nineteenth century and beyond), and in much greater numbers than is generally supposed. To be sure, they are not always easy to use: particular collections of budgets are rarely representative, seldom directly comparable across sources. But the tools of the modern welfare analyst can mitigate these problems in historical contexts too. There is thus a real opportunity to give a decisive boost to historical research on both the microeconomics of households and the macro dynamics of welfare; this is the case we make in this article. In Sections 2 and 3, we describe modern and historical household budgets, respectively, with a very brief review of their previous use in economic history. The focus then narrows to our own area of interest,
the long-run evolution of living standards, poverty, and inequality. In Section 4 we review existing historical research on these issues. Section 5 addresses technical problems arising in work with historical family budgets and in welfare comparisons over time and across countries. We outline our new Historical Household Budgets initiative in Section 6, and present some preliminary results for the Italian case in Section 7. We conclude our manifesto with a call to arms in Section 8.

2 In praise of modern household budgets

In current usage, a household budget is an accounting schedule that sets out the incomes and expenditures of a household with reference to a given period of time. Many people keep household accounts, often in an informal way, in order to keep an eye on the family’s incomes and expenditures. Besides being a domestic tool in citizens’ private lives, household budgets play a fundamental role in modern economics and statistics (Deaton, 1997). Each year, millions of individuals are interviewed the world over by an army of interviewers sent by national statistical offices, independent research institutes, and other organizations. In many countries, participation in household budget surveys is mandatory by law. What are all these data on household budgets for, and who uses them?

A convenient starting point is to note that household budgets, in contrast to other statistics, let us see how families respond to the economic environment. Wage data or unemployment figures might show us that, say, industrial workers’ living standards are under pressure. But this tells us little about how their families are affected or how they react. They might switch consumption from expensive meat to cheap starches, cut health care expenditures, or produce more at home for their own consumption; they might keep children out of school and put them to work; a stay-at-home spouse might take up market employment; the family might move to a smaller house, take in lodgers, or spend less on fuel; the head of household may work more days per year by taking on a secondary employment; they might maintain consumption by drawing down family

4 Households and families are the basic units of analysis here; they are not the same thing. A household is composed of one or more people who share the same dwelling, and not all households contain families. In turn, a family is defined as those members of the household who are related, to a specified degree, through blood, adoption or marriage. Currently, no single definition has general use around the world.
savings or mobilising community resources. Wage data tell us none of this. Even income data would fail to tell us how and why annual incomes were maintained despite falling wages. It is only when we know about consumption patterns, wealth and savings, housing, labour market participation of family members, and so on that we can fully understand a) the level and dynamics of family welfare, and b) the strategies families use to cope with economic adversity.

For the academic community, international organizations, and governments, household budgets are an essential input for economic analyses. Household incomes and/or expenditures are used to measure economic welfare. A household’s, or an individual’s, standard of living encompasses health status, education, freedom, access to basic services; it is more than just money. Still, if we seek one indicator, then consumption may be the best available approximation to utility, which is the starting point for how economists think about household behaviour and about welfare (Deaton and Grosh, 1998). Household budgets are therefore essential for analyzing the distribution of welfare and understanding its determinants.

More prosaically, household budgets are a foundation for estimates of private consumption, the most important component (55-60%) of an expenditure-based calculation of GDP (Lequiller and Blades, 2006: ch. 6). In many countries, household budgets are an important input into construction of the national income and product accounts. In other cases, they serve as a useful check on the plausibility of the official figures. Household budgets are the key to official estimates of inflation as well, for it is the consumption patterns that they reveal which provide the weights needed to compute a consumer price index, or alternatively a spatial price index measuring geographic variation in the cost of living (ILO, 2004). Detailed information on the consumption patterns of poor families is indispensable in identifying the minimum expenditure required to reach a socially defined acceptable standard of living, in other words an absolute poverty line.

Household budget data are similarly important in policy evaluation. They are required to calibrate the models used in microsimulation studies of, for example, how the labour supply of parents changes in response to changes in the tax-benefit system (Figari, 5 Of course the composition of household expenditure and its development over time interest the private sector too. They reveal the evolution of consumer tastes and preferences, i.e. the market that producers face, and are the key to producing accurate forecasts of demand (Thomas, 1987).
Paulus and Sutherland, 2014). The most important international agencies – from the United Nations to the World Bank – provide technical assistance to countries requesting it in order to design and implement surveys on household budgets: a key element for establishing development plans and for gaining access to international financing (Deaton, 1997).

As hinted earlier, an ideal conception of the standard of living would comprise more than just current consumption levels. It is disaggregated, family-level data that allow analysts to go further and adopt a multidimensional approach to welfare. Composite indices are a popular implementation of this approach (Anand and Sen, 1994), but have been shown to be deeply problematic (Fleurbaey, 2009; Amendola et al. 2016), and fail to take account of the distribution of their constituent indicators. Sufficiently comprehensive family-level data also allow us to shift the analysis from final achievements such as consumption to opportunities (Peragine, 2004). Another broadening of our conception of living standards is to consider vulnerability to poverty, which introduces the role of risk and the household’s capacity to mitigate it (Dercon, 2004). Again it is household-level data, ideally with a longitudinal dimension, that are needed to give empirical content to this idea.

What of the collection of household budget data? Modern surveys pay particular attention to two criteria: i) statistical representativeness, that is the units in the sample (e.g. households, individuals, etc.) are selected through a probabilistic sampling mechanism that ensures the data are representative of the target population (Levy and Lemeshow, 1999); and ii) consistency of interpersonal comparisons, that is variable definitions that ensure comparability across units (households or individuals) – and that are theoretically grounded (Deaton and Zaidi, 2002). Virtually all countries in the world conduct cross-section and longitudinal surveys on a regular basis for the purposes discussed in this section. On the theoretical side, most economists would not hesitate to argue that household budget surveys represent a “first best”. As with any source, they must be used with care and an awareness of their limitations, but they are the benchmark against which other data are compared.

---

6 This does not imply that they are perfect. Atkinson (2015: 48), among others, discusses their limitations. Meyer, Mok and Sullivan (2015) have argued that household surveys are in crisis, threatened by declining accuracy due to reduced cooperation of respondents. See also Burghauser et al. (2016).
3 Household budgets in history

Household budgets are not a recent invention; Cicero (106 BC – 43 BC) tells us that in ancient Rome citizens habitually kept two kinds of accounts: a sort of first entry (called adversaria) in which they hurriedly jotted down their everyday transactions, and a real book of expenditures (the codex accepti et expensi) in which the adversaria notes were properly organized on a monthly basis (Smith, 1873: 17). None of these documents has come down to us except in the form of indirect references in the letters of Cicero and other Latin writers, but they clearly testify as to how deeply rooted in the past the habit of family bookkeeping is (Chianese and Vecchi, 2016).

Nor do household budgets seem limited by narrow geographic boundaries. Western Europe and the North America have a well-known and strong tradition (Stigler, 1954; Bales, Bulmer and Sklar, 1991; Vecchi, 2016, 2011). North European countries have “data to die for” (see Jäntti, 2006; Saaritsa, forthcoming; Öberg, forthcoming). In the United States, studies on the living standards of the population were carried out on a large scale well before the advent of modern surveys: Eaton and Harrison (1930) listed 2,775 surveys between 1907 and 1908 (Stapleford, 2009), while earlier initiatives have been discussed in Converse (1987) and Roberts (forthcoming) documents the many sources available for other so-called western offshoots (particularly Canada, Australia and New Zealand).

Moving east, we find in Russia perhaps the most impressive trove of historical household budgets. These derive from the extraordinary efforts of local government (zemstvo, pl. -a) statisticians, who collected hundreds of thousands of peasant household budgets in the late nineteenth and early twentieth centuries (Seneta, 1985; Darrow, 2000). In carrying out tax assessments, zemstvo statisticians were also tasked with collecting data suitable to the development of informed policies for rural development. Fedor Andreevich Shcherbina (1849-1936), head of the Statistical Bureau of the Voronezh Zemstvo, near the northeast border of modern Ukraine, believed that only a household-by-household examination could do justice to the heterogeneity of Russia’s rural economy. Shcherbina’s studies, which were emulated by his zemstvo colleagues across the Empire, were used intensively by Lenin (Kotz and Seneta, 1990), and more

7 Relatively little is known about Latin American countries. For Mexico, López-Alonso (2012: 63-64) maintains that there are insufficient household budget data for the period prior to 1957.
recently by Field (1989) and Lindert and Nafziger (2012). Much of the original microdata is now available – already digitized – to interested scholars.

Proceeding further east, historical family budgets are found in China (Fang, Wailes and Cramer, 1998), India, as well as in Japan. In the latter country, their appearance dates back to the Edo period (Ogura, 1982), though most sources refer to the Meiji and Taishō periods. It is worth noting that in Japan, perhaps more than in other countries, household budgets were rooted in the national culture and formed a part of social habits well before the modern survey era. Signs of their presence in the cultural life of the populace are everywhere. In the magazine Fujin no Tomo (The Women’s Friend, est. 1911), Hani Motoko (1873-1957) published a set of household finance ledgers (kakeibo), encouraging women to rationalize management of the household (Komori, 2007). The diffusion and success of the kakeibo on a national scale was astonishing – “Readers eagerly welcomed those articles as ‘helpful’, ‘convenient’ and ‘interesting’, and the magazine became one of the most popular magazines among Japanese women” (Hiroko, 2004: 55-56). It is no accident that large-scale surveys on household income and expenditures were promoted by Statistics Bureau of Japan as early as 1926 (Smitka, 1998).

Finally, Africa. In 2010, Shanta Devarajan, chief economist for the African region at the World Bank, gave a keynote speech deploring “Africa’s statistical tragedy”, that is, a chronic “weak capacity in countries to collect, manage, and disseminate data” (Devarajan, 2013: S12). Does Devarajan’s critique apply equally to African history? From this angle, Africa remains a largely unexplored continent. With reference to the 20th and 21st centuries Jerven (2015) has taken up the challenge by investigating the history of measurement for selected countries. Serra (2014) has focused on Ghana, Marivoet and De Herdt (2015) on the Democratic Republic of Congo, Davie (2015) on South Africa. Much more is available (ILO, 1926, 1949, 1961; Woodbury 1940), even if currently filed under “research agenda for the future”. We will return to this issue in Section 7.

---

8 Kakeibo have survived the test of time and reached global fame. When writing this article, Italian translations of kakeibo could be seen on display among the bestsellers at the main bookshop of the Rome airport. Internet bookshops too are well stocked.
Table 1. Historical household budget stock estimates, 1850-1965

<table>
<thead>
<tr>
<th>Country</th>
<th>Households frequency</th>
<th>Households percentage</th>
<th>Country</th>
<th>Households frequency</th>
<th>Households percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>375,828</td>
<td>20.5</td>
<td>Sweden</td>
<td>8,251</td>
<td>0.5</td>
</tr>
<tr>
<td>Germany</td>
<td>349,036</td>
<td>19.0</td>
<td>Canada</td>
<td>7,814</td>
<td>0.4</td>
</tr>
<tr>
<td>United States</td>
<td>298,588</td>
<td>16.3</td>
<td>Ghana</td>
<td>6,958</td>
<td>0.4</td>
</tr>
<tr>
<td>Russia</td>
<td>179,754</td>
<td>9.8</td>
<td>Egypt</td>
<td>5,283</td>
<td>0.3</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>108,895</td>
<td>5.9</td>
<td>Singapore</td>
<td>5,225</td>
<td>0.3</td>
</tr>
<tr>
<td>India</td>
<td>56,993</td>
<td>3.1</td>
<td>Jamaica</td>
<td>4,119</td>
<td>0.2</td>
</tr>
<tr>
<td>Italy</td>
<td>52,888</td>
<td>2.9</td>
<td>Senegal</td>
<td>1,545</td>
<td>0.1</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>31,420</td>
<td>1.7</td>
<td>Alaska</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>Brazil</td>
<td>30,854</td>
<td>1.7</td>
<td>Europe</td>
<td>1,143,073</td>
<td>62.3</td>
</tr>
<tr>
<td>Argentina</td>
<td>27,656</td>
<td>1.5</td>
<td>America</td>
<td>422,685</td>
<td>23.0</td>
</tr>
<tr>
<td>Philippines</td>
<td>25,883</td>
<td>1.4</td>
<td>Asia</td>
<td>173,237</td>
<td>9.4</td>
</tr>
<tr>
<td>Japan</td>
<td>25,079</td>
<td>1.4</td>
<td>Africa</td>
<td>94,873</td>
<td>5.2</td>
</tr>
<tr>
<td>Spain</td>
<td>24,022</td>
<td>1.3</td>
<td>Oceania</td>
<td>2,219</td>
<td>0.1</td>
</tr>
<tr>
<td>Southern Rhodesia</td>
<td>24,012</td>
<td>1.3</td>
<td>World</td>
<td>1,836,087</td>
<td>100.0</td>
</tr>
<tr>
<td>China</td>
<td>20,559</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ceylon</td>
<td>14,518</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors’ preliminary estimates, mainly based on English-written published sources.

Table 1 reports estimates of available household budgets for the period prior to modern probabilistic surveys by country and macro region, based on a rapid reconnaissance of documented sources. A precious starting point is a 1935 US Department of Agriculture publication, *Studies of families living in the United States and other countries: an analysis of material and method*, which surveyed then-known sources of information on household living standards around the world from the second half of the nineteenth century to 1930. Roughly 1,500 printed works from 52 countries were listed, most of which appear to have household budgets (Williams and Zimmerman, 1935). Other compilations – such as *Bulletin de la Statistique Générale de la France*, various years, ILO (1926), Halbwachs (1913, 1933), Staehle (1934, 1935) and others – point to additional sources with thousands of household budgets.
Two comments on the figures in Table 1 are in order. First, the geographical distribution is highly unbalanced (three countries account for more than 50% of total observations; two thirds of all observations are for European countries). This is a consequence of our preliminary research strategy rather than the likely availability of actual data. Our research of has focused on countries for which the language barrier is not an obstacle, which implies that Arabic, Chinese, some Slavic, and many South Asian sources have not been even approached. This leads to the second remark, which is that Table 1 is likely to show only the tip of the iceberg: for most countries we expect that historical sources written in local languages outnumber sources in English and other languages in which our compilations authors were proficient. In the case of Italy, a back-of-the-envelope calculation suggests that the multiplier exceeds one hundred: for each budget known to authors of the compilations, more than 100 where found in Italian libraries, archives and secondary sources. The upshot of Table 1 is that huge numbers of historical household budgets are available around the world.

Despite the relative abundance of source material, historical studies of household were relatively few until relatively recently. But that is changing. Sara Horrell, Jane Humphries and Deborah Oxley have pushed the approach well back in time, and also pushed the limits of what we can learn. Following in their footsteps, other scholars have used historical household budgets to study a range of themes. These include diet and nutrition (Vecchi and Coppola, 2006; Logan, 2006, 2009; Gazeley and Horrell, 2013; Gazeley and Newell, 2015; Lundh, 2013), inequality and poverty (Rossi et al. 2001, Amendola and Vecchi, 2016), intra-household dynamics (Horrell and Oxley, 2013; Saaritsa and Kaihovaara, 2016; Scott, Walker and Miskell, 2015; Guyer, 1980), labor force participation (Baines and Johnson, 1999), child labor (Moehling, 2001, 2005), consumption behavior (Scott and Walker, 2012; Lilja and Bäklund, 2013), agriculture and home-production (Federico, 1986, 1991), economies of scale and child well-being (Hatton and Martin, 2010; Logan, 2011), and informal transfers of cash and goods between households (Saaritsa, 2008, 2011).

Overall, the potential of household budgets for economic history has been exploited to only a limited extent. Why? Because household budgets, as found in historical sources, are very different from modern household budgets and historical household budgets

---

9 See, for example, Horrell and Humphries (1992; 1995; 1997), and Horrell and Oxley (1999; 2000).
cannot be treated as if they were modern household budgets: a collection of household budgets, however large it may be, is not a representative sample and cannot unproblematically be used for statistical inference about living conditions in the population. Nor are historical data harmonized; the heterogeneity of the sources makes the budgets too diverse to allow for consistent comparisons (Lanjouw and Lanjouw, 2001). Awareness of these limitations helps explain why economic historians have been reluctant to use household budgets for the pre-statistical era. They believed, in short, that household budgets were too scarce and sparse, too problematic and unreliable to permit generalizations and to address big questions. Skepticism prevailed.

In the remainder of this article we narrow our focus to a specific theme among the many that household budgets can illuminate – the long run evolution and distribution of living standards. This is a field where a number of prominent figures have concentrated their efforts (Steckel and Floud, 1997; Fogel, 2004; Allen et al. 2005, van Zanden et al. 2014; Gordon, 2016). We illustrate the potential contribution of historical household budgets, and present arguments that ought to win over the skeptics.

4 Inequality and poverty in the long run: the state of the art

While modern analysts can rely on household budget surveys, the perceived lack of suitable data has been a major obstacle for economic historians. Several lines of inquiry have emerged, even if we restrict our attention to monetary indicators and set to one side approaches such as anthropometric history. We distinguish three schools and dub them the Eclectics, the Heroics, and the Fiscalists. The Eclectics have discovered ingenious sources and methods (e.g., Soltow, 1968; Williamson and Lindert, 1980; Williamson, 1985; Lindert, 2000). A common denominator is the use of proxies for living standards, such as occupational pay ratios, the window tax collected under the Inhabited House Duty in Britain, and probate records (van Zanden et al. 2014). Reactions have not always been enthusiastic (Feinstein, 1988). Moreover, the distance from the benchmark, the methods one would employ with modern household budget surveys, is ample: proxy variables perform

---

10 While the names Eclectics and Fiscalists are self-describing, Heroics perhaps deserve an explanation. The idea is to evoke antiquity, about which some in the school have written, and the spirit of enterprise in calculating inequality in a remote period from rather sketchy data.
poorly in terms of both population coverage and theoretical consistency of the welfare indicator.

That said, it is difficult to overstate the importance and influence of this first generation of studies. They can be thought of as the second pillar of a bridge taking us from a lack of interest in and knowledge about the historical distribution of income to an empirically-grounded understanding of the long-run dynamics of inequality. The first pillar is represented by Kuznets’s (1955) classic. As Kuznets tried to squeeze the most out of his scant data, so did Lindert, Williamson and other scholars with their eclectic conjectures. They succeeded in writing highly innovative pages of economic history, opening up a new research field that is still flourishing today. Lindert and Williamson’s (2016) latest book is the best evidence in support of the well-deserved fame of this school.

The *Heroics* have explored the use of “social tables” for the pre-industrial era: tabulations by contemporaries of average family incomes in different social strata alongside population shares of these groups. The allure of social tables is hard to resist, because they allow the analyst to roam widely in the past, from the Roman Empire in the year 14 (Scheidel and Friesen, 2009) or Byzantium in year 1000, to Moghul India in 1750 or the Kingdom of Naples in 1811 (Milanovic, Lindert and Williamson, 2011; Milanovic, 2011). On the basis of social tables van Zanden (1999) has estimated inequality in European cities from 1500s onward, Hoffman et al. (2002) in European countries since 1500, and Lindert and Williamson (2012) in the US over its first century (1774-1860).

In parallel with the rediscovery of social tables, more refined statistical techniques have been developed to analyze them. The tables report mean incomes only, offering no information on within-group variation. Modalsli (2015) has recently suggested a procedure for taking this into account when calculating inequality. Other efforts have been directed at identifying the best choice for modeling income distributions, that is, of interpolating income distributions. Interpolation, far from being an old-fashioned topic, is highly relevant to historical studies (Atkinson, 2007: 39). Already a lengthy menu of functional forms has been considered, including the lognormal, gamma, Pareto,

---

11 Lindert (2000: 174) writes that “(f)or Britain before 1914, our best guesses are necessarily eclectic. There is little choice but to weave an archival quilt of indirect clues on income inequality.”

12 ‘Best’ here mainly denotes superior empirical goodness-of-fit.
Weibull, Dagum, Singh-Maddala, Beta and Generalized Beta of the second kind distributions (Jenkins, 2009). New studies continue to expand this list, increasing the robustness and credibility of social table–based estimates (Hajargasht and Griffiths, 2013; Chotikapanich et al. 2013).

The Fiscalists, the school of Piketty and collaborators, rely on tax records. Adopting the pioneering methods of Kuznets (1953) and Atkinson and Harrison (1978), the fiscalists have succeeded in computing the shares of top incomes in the total, often on a yearly basis, for a number of countries. Common methodological features of the fiscal approach are the use of i) income tax data, ii) national accounts, and iii) Pareto interpolation (that is, the assumption that the unobserved incomes below those at the top follow a Pareto distribution). The relative abundance of ingredients i) and ii) combined with the relentless activity of the school’s economists is the source of the recognition it has received.

Results have been rich and have stimulated an intense debate both within academia and among the public. An impressive two-volume work edited by Atkinson and Piketty (2007, 2010), as well as a host of other publications – too many to be catalogued here – have produced a large volume of data (Alvaredo et al., 2013). Since 2011 the World Top Incomes Database (WTID) has made many findings publicly available for download. Among the important results generated thus far is the empirical validation of the Kuznets’s curve.

The work of the three schools is inspiring. But there remains room for improvement. None of the sources used by the Eclectics meets the standard of modern analysis, neither conceptually (the proxies for living standards do not have the properties desirable for a comprehensive welfare indicator – see Deaton and Zaidi, 2002) nor empirically (the sources employed being extremely heterogeneous). Nor do these sources allow the analysts to tackle the multidimensional nature of living standards.

True, a number of difficulties have been overcome by the Fiscalists, but many limitations remain. First, most top income share series start around the time of the First World War, so time coverage is limited. Second, inequality measures based on tax records are silent about what happens to the bottom 90% of the population, or even the bottom 99%. This is a most serious limitation – modern analysts would despair in the absence of 90-99% of the target population, and would mount a critique without mercy.
The nature of the problem is not too different from what occurs in modern surveys, when investigators are faced with low response rates. There are remedies for dealing with nonresponse, but they are effective only up to a point: beyond a certain threshold, the entire survey becomes worthless (Lohr, 2009: ch. 8). Third, by failing to reconstruct the entire income distribution, the WTID does not allow us to estimate the long-run trend in poverty. This is worth emphasizing given the intellectual interest in the extent to which economic growth is inclusive, and more specifically pro-poor. Fourth, WTID considers only the income dimension of welfare and does not allow for multidimensional poverty or inequality measures (Bourguignon and Chakravarty, 2003, Alkire and Foster, 2011; Aaberge and Brandolini, 2014).

The limitations above discussed are of major concern to analysts interested in the long-run evolution of living standards. A fourth school – still in search of a name – has been founded on the idea of sending a modern welfare analyst back in time. Taking advantage of i) knowledge and expertise in dealing with modern household budget data and ii) the abundance of historical household budgets documented in Section 3, the modern analyst working in the past would not seek new or more ingenious proxies for living standards, but would construct a family budget based welfare indicator following the best current practice. In so doing she would have a good chance of approaching the “first best” solution discussed in Section 2. If anything like this were possible, it would be a real advance in our understanding. Few economic historians have taken this path so far, but interest is growing. Can we not send the welfare analyst back in time?

5 Working with historical household budgets – a user guide

Given the potential of historical household budgets, we need to discuss the ways they differ from their modern counterparts. The key problems for which we need solutions are those discussed in Section 2, namely i) non-representativeness (a historical collection is not a statistical sample), and ii) non-standardization (historical sources vary widely across many dimensions). Both issues pose serious conceptual and technical problems; in this section we discuss how the time-traveling welfare analyst can handle them, applying lessons learned by development economists immersed in a daily struggle with troublesome data.
Non-representativeness. The lack of an underlying probabilistic survey design means that no collection of historical budgets, no matter how large, can be assumed be a statistical sample representative of the population. Some strata will be under-represented, others unrepresented altogether in the data. Post-stratification (stratification after collection of the sample) offers a viable solution to this problem (Cochran, 1977: Ch. 5A; Holt and Smith, 1979). The idea is simple: we weight the budgets in our historical collection so as to make its composition resemble the population. Where to obtain appropriate weights? We can derive them from a population census in a year close to that of our budgets. The census gives the frequencies of different types of households: by region, by family composition and ages, by occupation of the head of household, and so on. These frequencies are the weights or “expansion factors” to be applied to the corresponding budgets in our historical collection. The weighted budget collection can then be treated as if it were a probabilistic sample.

Post-stratification is widely used in statistics, but is not a guarantee of success (Lohr, 2009: 114). A necessary condition is that the initial collection of budgets be sufficiently numerous and representative of the target population. To clarify, with a collection of only urban households there is no redemption, no way to magically produce nationally representative estimates without some coverage of the rural sector. Assuming the data are not affected by major problems of “non-coverage”, the question becomes: how reliable is post-stratification, when applied to historical household budgets? The evidence for Italy is that post-stratified statistics based on historical budget collections agree closely with statistics based on the national accounts such as GDP per capita, which is a non-trivial result (Ravallion, 2001; Deaton, 2005).

Figure 1 illustrates these issues in the case of the United Kingdom in the early 1900s. The top panel shows the consequences of non-coverage; it plots the empirical cumulative distributions of household income per capita in two samples. The Board of Trade (BoT) sample comprises urban households only; the supplemental sample includes both urban and rural households. It is immediately apparent that the BoT

---

13 “Non-coverage” – the complete omission from the sampling mechanism of some segments of the population, for instance rural households in a particular region, forces the researcher to turn to external, non-sample sources of information. Little and Rubin (2002) review a number of statistical techniques available to deal with non-coverage.
distribution lies everywhere to the right of (or below) the supplemental sample. If we take *any* level of household income, a smaller share of the BoT households are below it, a larger share above, compared with the supplemental sample. In technical terms, the BoT sample distribution first-order stochastically dominates the new sample (Atkinson, 1987). This first order stochastic dominance means that any poverty line adopted will yield a smaller incidence of poverty – additional evidence beyond what we know from documentary evidence that the BoT sample was not (intended to be) representative of the entire population.

The effect of combining different sources, even if these originate entirely independently and not as part of any coordinated survey process, is to mitigate the bias of any single sample, and the inclusion of rural households clearly represents an improvement here. But this is not sufficient to make the collection representative of the population of Edwardian Britain. A second problem is caused by the unbalanced distribution of the observations between rural and urban, or, within these categories, between regions or sectors of activity. This is where post-stratification fits in.
Figure 1. Household budgets in the United Kingdom, early 1900s

Source: A’Hearn, Di Battista and Vecchi (mimeo).

The lower panel of Figure 1 shows the effect of post-stratification based on two regions (North and South) and three sectors of household head employment (agriculture,
industry, and services). One can see how post-stratification changes the shape of the distribution, and with it poverty and inequality estimates.\textsuperscript{14} The vertical gap between the curves in Figure 1’s lower panel is quite substantial at many points: an apparently small distance translates into hundred of thousands of individuals. If we draw a poverty line around 6.8 p/person/day, for example, post-stratification implies 1.5 million additional individuals in poverty relative to the count that results from the raw, unweighted data.

Table 2. Household budgets in Italy, 1861-1961: selected sources

<table>
<thead>
<tr>
<th>Source type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large surveys (non-probabilistic)</td>
<td>Occasional investigations into social and economic conditions initiated by parliament or ministries, which could yield hundreds of HBs.</td>
</tr>
<tr>
<td>Government personnel files</td>
<td>Dossiers maintained for all employees, from the minister to the cleaning staff, typically contain enough information to form an HB.</td>
</tr>
<tr>
<td>Court records</td>
<td>Commercial law, particularly regarding bankruptcies, often required complete data on a debtor’s family financial circumstances.</td>
</tr>
<tr>
<td>Family monographs</td>
<td>There are hundreds examples of detailed studies in the tradition of Le Play.</td>
</tr>
<tr>
<td>Periodical press articles</td>
<td>Local newspapers often collected and published accurate, meticulously-documented HB data.</td>
</tr>
<tr>
<td>Hospital records</td>
<td>Charges for medical services were often proportional to family income, forcing hospitals to collect and record HB data.</td>
</tr>
<tr>
<td>Insane asylum records</td>
<td>Mental illness could be grounds for confinement in an asylum or a government subsidy for home care. Either required a statement of HH resources.</td>
</tr>
<tr>
<td>Occasional local censuses</td>
<td>Occasional local censuses, e.g. a census of housing in a booming city, in which families were required to report their accounts.</td>
</tr>
<tr>
<td>Foreign sources</td>
<td>Studies by foreign powers, e.g. US Dept. of Commerce or UK Board of Trade, interested in labour costs of competitors.</td>
</tr>
<tr>
<td>Family archives</td>
<td>Upper middle class and aristocratic families kept accounts, often professionally.</td>
</tr>
<tr>
<td>Firm personnel files</td>
<td>Large firms, e.g. FIAT or Peroni, monitored workers’ living standards by regularly collecting information on income and family circumstances.</td>
</tr>
<tr>
<td>“Typical” and hypothetical HHs</td>
<td>Compiled for particular groups (textile workers in a region, say) by contemporaries or retrospectively on the basis of separate data on wages, the cost of living, family composition, etc.</td>
</tr>
<tr>
<td>Social assistance requests</td>
<td>Applications for invalidity pensions, subsidised housing, and the like typically required documentation of family circumstances.</td>
</tr>
</tbody>
</table>

\textsuperscript{14} The vertical jumps that can be observed in both curves are a consequence of a number of factors, which include the uneven distribution of observations within and between strata.
Harmonization. The harvesting of historical household budgets implies recourse to heterogeneous sources. In the case of Italy, Somogyi (1973) called household budgets a “kaleidoscopic mosaic,” suggesting intricate detail and changing perspectives. But the metaphor also reminds us of the fragmented and varied nature of the data. In historical budget inquiries varying and imperfect coverage was the rule, not the exception. Sample sizes varied from just one (!) – in the cases of some family monographs or “typical” family budgets compiled by informed contemporaries – to a thousand and more. 15 Table 2 illustrates the range of sources discovered and exploited thus far for Italy.

Constructing a standardized dataset is not a simple undertaking. Practically, it implies that each record must be processed to conform to a uniform scheme (Browning, et al., 2003) but historical sources vary widely: they have different reference periods (from a single day up to an entire year), employ different methods of capturing data (diary vs. recall), and report information with different levels of detail (at one extreme many sources provide only total household income; at the other are sources with 50 and more categories of expenditure). All this variety implies that the data – as reported in the original sources – are not comparable (Beegle et. al 2014).

Two further issues are worth mentioning here. The first concerns grouped data. Precious information, typically from mid-twentieth century sources, is frequently presented in the form of histograms, or grouped data: separate tabulations giving the number of households in different income categories, in different expenditure categories, in different regions of residence, occupations or sectors of activity of the household head, etc. Such empirical distributions allow us to calculate some measures of interest directly, e.g. mean income, but they do not allow us to calculate, say, the share of rural farming families that fall below a poverty line. Fortunately, new techniques are being developed that aim to extract synthetic household-level data from such tabulations (Shorrocks and Wan, 2008), and ongoing research in the area promises further improvements in our ability to exploit historical data on household budgets. The second issue relates to purchasing power parity exchange rates.

15 The most famous example is probably Frédéric Le Play’s (1806-1882) family monographs. See Lorry (2000).
5.1 Purchasing power parity (PPP) exchange rates.

Market exchange rates are the obvious candidate for converting local currency measures to a common basis for comparison. Obvious, perhaps, but not obviously correct, since market rates do not always reflect the actual purchasing power of currencies. Economists today rely on PPP exchange rates, but reliable PPP standards, alas, are not available for the period before 1950 (Deaton, 2010; Deaton and Heston, 2010).

A useful starting point for the discussion is to set out the issues algebraically. We begin with household expenditure on goods and services. Let us denote by $E_j^h$ the total expenditure of household $h$ in country $j$:

\[
E_j^h = \sum_{g=1}^{G_j} x_{j,g}^h
\]

In Eq. (1) $x_{j,g}^h$ is household expenditure on a single good $g$, and the summation notation indicates that we add up expenditures on all items in the set $G$. Typically, different countries have different consumption habits, hence different, country-specific bundles of goods and services, $G_j$. These consumption baskets varied enormously across countries and over time: imagine Ghana in 1850 and Sweden today. The relevance of this point should not be underestimated; if the consumption bundles of two countries show little or no overlap, the basis for meaningful cost-of-living comparisons between countries becomes difficult, if not impossible.

Suppose we wish to compare expenditure across two households, say $h$ and $h'$, in two countries, say $i$ and $j$. Each of $E_i^h$ and $E_j^{h'}$ is expressed in the relevant local currency. The simplest conversion uses the market exchange rate $XR_{i,j}$ between the two currencies:

\[
NE_j^h = \frac{E_j^h}{XR_{i,j}}; \quad NE_i^{h'} = \frac{E_i^{h'}}{XR_{i,i}}
\]

where $NE_j^h$ and $NE_i^{h'}$ are the expenditures of the two households expressed in a common currency, the numeraire. In this case, country $i$'s currency is the numeraire and $XR_{i,j}$ is the price of currency $i$ in terms of currency $j$. If the dollar is the numeraire, $XR_{i,j}$ has units like £0.7/$ and $XR_{i,i}$ is $1/$. Thus $NE_i^{h'} = E_i^{h'}$. 

The problem with using $NE_j$ and $NE_i$ for welfare comparisons is that the market exchange rate used in the conversion may not reflect the actual purchasing power of the currencies in their home countries (Rogoff, 1996). It is not based on the prices faced by households in their respective countries, nor on their different consumption patterns. To better understand the issues, we define a spatial price index, or SPI. The SPI can be calculated as the relative cost of buying the same basket of goods in two countries, using the market exchange rate to express both in terms of the numeraire (Diewert, 2003).

The cost of the basket in local currency is:

$$E_j = \sum_{g=1}^{G} p_{j,g} \cdot q_g$$

where $g = 1 \ldots G$ indexes the consumption items in the basket, the $q_g$’s are the weights of each item in the common basket, and the $p_{j,g}$’s are the country-specific prices of each commodity in the local currency. The formula for $E_i$ is analogous, with prices $p_{i,g}$. As before we convert local currency measures to the numeraire using the market exchange rate:

$$NE_j = \frac{E_j}{XR_{i,j}}$$

with $XR_{i,j}$ again equal to one by definition. Finally we have

(2) \hspace{1cm} SPI_j = \frac{NE_j}{NE_i}$$

The SPI gives us the relative price level in the two countries. When SPI exceeds one, country $j$’s prices are “too high” relative to country $i$. Alternatively, its currency is overvalued relative to purchasing power parity. The hypothetical exchange rate that would preserve the purchasing power of our money when changing currencies would be:16

(3) \hspace{1cm} PPP_{i,j} = XR_{i,j} \times SPI_j$$

---

16 An algebraically equivalent formulation is $PPP_{i,j} = \frac{E_j}{E_i}$. 
We can use this PPP standard to convert the local nominal expenditures of our two households to real expenditures expressed in terms of the numeraire.

\[
RE_j^h = \frac{NE_j^h}{PPP_{i,j}}
\]

\[
RE_i^{hv} = \frac{NE_i^{hv}}{PPP_{i,i}} = E_i^{hv}
\]

PPP estimation thus requires two ingredients, namely market exchange rates and spatial price indices (Deaton and Muellbauer, 1980: Ch. 7; World Bank, 2013). The former ingredient is not problematic: economic historians have been studying market exchange rates for a long time. Regarding the SPI, its estimation in a historical context is more complex.

Construction of an SPI requires i) a set of homogeneous consumption goods and services \((g=1...G)\) common to country \(i\) and country \(j\); ii) market prices \(p_{j,g}\) and \(p_{i,g}\) of these goods and services, and iii) the corresponding expenditure budget shares \(q_g\) used to weight items in the basket. In historical settings, information about prices is fragmentary, often restricted to a few tradable goods (Allen, 2001). But for this subset of goods, precisely because they are tradable, the law of one price is likely to hold, in other words \(p_{j,g} = XR_{i,j}p_{i,g}\) (otherwise there would be arbitrage opportunities, abstracting from transaction costs). It is non-tradable goods – and services – that motivate our interest in PPP in the first place. Even more difficult is the reconstruction, with the required detail, of the average consumption pattern of the households. In practice, measurement errors in i)-iii) can be large. In the absence of high quality ingredients, a reasonable and defensible alternative may be to stick to the simpler recipe of market exchange rates, using directly \(NE_j^h\) as a welfare indicator for international comparisons. We will return to this in Section 7, when comparing Italy with the United Kingdom in the early 1900s.
6 The Historical Household Budget (HHB) Project

The notion of combining tools from welfare economics and statistics with historical (and modern) household budgets is more than just an idea. The authors of this paper have launched an ambitious project – the Historical Household Budgets (HHB) Project – to promote research into the evolution of living standards, within and between countries, over the last two to three centuries.\(^\text{17}\) The project is interdisciplinary, long-term, and unique in the way it has been envisaged.\(^\text{18}\) Here we offer a brief description.

HHB is based on a collaborative, open-source vision of research, and aims to provide a publicly-available set of tools for working with household budgets that brings together interested scholars from around the world. The Project is currently developing a web-based infrastructure to host an HHB database. Given the abundance of data documented earlier, and our ongoing bibliographic search for additional sources, we predict that the database will eventually include several million household-level records, as well as several terabytes of digital reproductions of source documents. But the database is about more than just storage; the HHB web platform will allow researchers to enter, query, and analyse household budget data in such a way as to permit seamless links between countries and over time. Harmonisation is what this system aims at.

The HHB database (Table 3) comprises ca. 500 variables including monetary measures of the standard of living (consumption, income, wealth, but also wages and retail prices), education and health, anthropometric measures, fertility, employment and migration, housing, agriculture, access to credit, and exposure to shocks.\(^\text{19}\) The design is based on the World Bank’s Living Standard Measurement Study (Grosh and Glewwe, 2000) and the Luxembourg Income Study’s efforts to harmonize micro-datasets from

\(^{17}\) In fact, the period covered by the project covers household budgets that extend back to the 15\(^\text{th}\) century – see Section 7. Sources for that era are sporadic and, presumably, rare.

\(^{18}\) Following Berry, Bourguignon and Morrison (1983), and Bourguignon and Morrison (2002), other projects have been developed, such as the University of Texas Inequality Project (http://utip.lbj.utexas.edu), and the “Global Consumption and Income Project” (http://gcip.info), described in Jayadev, Lahoti and Reddy (2015). These projects (as all other projects we are aware of) do not use historical microdata, that is household-level data before 1950, nor do they have a proper historical dimension. Much of their focus is instead on methodological issues. More importantly, none of the existing projects collects, harmonises and shares the microdata with users. None embraces the open-source vision that characterizes the HHB project.

\(^{19}\) Truth in advertising requires a disclaimer. The database has 500 variables to encompass the information reported across the full range of sources, but no individual source has so many items. Some sources in the database really are quite rich, though, e.g. Peruzzi’s monograph on a family of Tuscan sharecroppers in 1857. This study, conducted according to Le Play’s method, yields observations on 277 variables.
upper- and middle-income countries (Smeeding, Schmaus and Allegreza, 1985). Such ex-post data harmonization is a task that ranks high on the agenda of many international agencies, and HHB’s partnership links with LIS and the World Bank ensure that its historical budgets will continue to be linkable with modern budgets.\textsuperscript{20}

Given the tremendous variation across countries and over time, we need a scheme flexible enough to accommodate diversity, a scheme both international and historical; at the same time, we need this scheme to link to existing classifications. Variables must be defined on a comparable basis, caloric content of food quantities must be calculated, sometimes-archaic local units of measure must be automatically converted to a standard reference, and so on. To cite but one of hundreds of examples, consider the staple carbohydrate of many family diets, rice. There are many (more than 40,000!) varieties of rice, which differ in price and nutritional content; we know that “rice” in New York in 1895 is not the same as “rice” in Bombay in 1923. The HHB database goes beyond the UN’s “COICOP” classification system, which does not distinguish rice by type, and prompts the user to chose one of 44 basic types that can be identified from historical sources.\textsuperscript{21}

\textsuperscript{20} The construction of harmonized databases for exploring global trends in living standards is flourishing (Burkhauser and Lillard, 2005). In 2015, a special issue of the Journal of Income Inequality reviewed eight large-scale cross-national inequality databases – see Ferreira, Lustig and Teles (2015).

\textsuperscript{21} COICOP stands for Classification of Individual Consumption according to Purpose. Rice has the code 01.1.1 “rice in all forms”.

<table>
<thead>
<tr>
<th>Section</th>
<th>Variables</th>
<th>Section</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Metadata</td>
<td>51</td>
<td>10. Credit</td>
<td>18</td>
</tr>
<tr>
<td>Individual weights, data quality, comments</td>
<td></td>
<td>Household access to credit and financial</td>
<td></td>
</tr>
<tr>
<td>by the dataset managers and other</td>
<td></td>
<td>services.</td>
<td></td>
</tr>
<tr>
<td>technical information.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Household Roster</td>
<td>44</td>
<td>11. Income</td>
<td>8</td>
</tr>
<tr>
<td>Household composition and socio-demographics.</td>
<td></td>
<td>Household income by source.</td>
<td></td>
</tr>
<tr>
<td>Anthropometrics, health and access to</td>
<td></td>
<td>Information on jobs (sector, duration,</td>
<td></td>
</tr>
<tr>
<td>sanitation, by person.</td>
<td></td>
<td>wage, ...), by job.</td>
<td></td>
</tr>
<tr>
<td>4. Education</td>
<td>14</td>
<td>13. Agriculture</td>
<td>87</td>
</tr>
<tr>
<td>Educational achievements and school</td>
<td></td>
<td>Agricultural activities of household</td>
<td></td>
</tr>
<tr>
<td>participation, by person.</td>
<td></td>
<td>members in detail.</td>
<td></td>
</tr>
<tr>
<td>5. Migration</td>
<td>18</td>
<td>14. Non-Agricultural Enterprise</td>
<td>47</td>
</tr>
<tr>
<td>Migration (internal and international) history,</td>
<td></td>
<td>Non-Agricultural household enterprises</td>
<td></td>
</tr>
<tr>
<td>by person.</td>
<td></td>
<td>in detail.</td>
<td></td>
</tr>
<tr>
<td>6. Expenditure</td>
<td>13</td>
<td>15. Time Use</td>
<td>36</td>
</tr>
<tr>
<td>Household expenditures collected at the</td>
<td></td>
<td>Time allocation of household members, by</td>
<td></td>
</tr>
<tr>
<td>COICOP level.</td>
<td></td>
<td>activity.</td>
<td></td>
</tr>
<tr>
<td>7. Durables</td>
<td>12</td>
<td>16. Shocks</td>
<td>16</td>
</tr>
<tr>
<td>Durable goods owned by the household, by</td>
<td></td>
<td>Shocks impacting the household in the</td>
<td></td>
</tr>
<tr>
<td>item.</td>
<td></td>
<td>past year, by event.</td>
<td></td>
</tr>
<tr>
<td>8. Wealth</td>
<td>8</td>
<td>17. Community Data</td>
<td>14</td>
</tr>
<tr>
<td>Household wealth by asset type.</td>
<td></td>
<td>Facilities available in the community, by</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>facility.</td>
<td></td>
</tr>
<tr>
<td>9. Housing</td>
<td>46</td>
<td>18. Subjective wellbeing</td>
<td>6</td>
</tr>
<tr>
<td>Housing conditions of the household.</td>
<td></td>
<td>*Qualitative assessment of household</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>wellbeing and life satisfaction.</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>502</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7 On the job with the time-travelling welfare analyst

The approach described in earlier sections has been applied with some success, at least in the Italian case. Here we give a flavour of the results generated by this research programme, continuing to focus on the inter-household distribution of welfare. Linking historic records of the sort set out in Table 2 with more recent survey data collected by the Bank of Italy and the Italian Statistical Institute (Istat), Amendola and Vecchi (2016) have produced estimates of the distribution of household income per capita that span the full 150 years of Italy’s history as a unified country. Figure 2 plots the estimated distributions for selected years, illustrating the progressive enrichment of Italian families. More subtly, the shape of the distribution also changes over time. Note, for example, how below 2,500 euros per person, the 1981 distribution lies beneath the 2014 distribution: despite considerable growth in mean income, a greater share of households lay below this threshold, in poverty, in the later year.

Figure 2. The distribution of household income: Italy, 1861-2014

Source: authors’ elaboration.
The changing shape of the curves in Figure 2 reflects differential income growth for households in different parts of the distribution. This dynamic can be brought out in a more fine-grained way by means of growth incidence curves (GICs). GICs plot the growth rate of real income against initial level of real income, or income percentile (Ravallion and Chen, 2003). A downward sloping GIC indicates that the lowest incomes grow fastest, contributing to an equalization of the distribution; an upward sloping curve indicates the opposite. Figure 3 plots GICs for four sub-periods. In each panel, the broken horizontal line indicates mean growth across all households over the relevant period. The upper left panel depicts the growth experience of Italy’s first fifty years. Armed with only a top 10% income share, we would conclude that Italy had grown substantially more equal in this period of liberal democracy. A decline in the Gini coefficient (from 50.4 to 46.0) suggests the same (Vecchi, 2011: p. 430). But the GIC reveals that reality was more complicated, with the poorest families also losing out in relative terms. Moving to the upper right panel, we see a generally upward-sloping, inequality-generating GIC for the period spanning the early years of fascist rule and the onset of the Great Depression. The vertical scale is different here, such that lower half of the distribution experienced not only a relative but also an absolute decline in household income per capita. Interestingly, the GIC shows that the top 5% shared this decline. The net result in terms of the Gini index is constant inequality over the 1920s.

The lower panels of Figure 3 divide Republican Italy into two periods with a watershed near 1992’s currency crisis, in which the country was forced to devalue the lira and exit the European Exchange Rate mechanism. In the earlier period growth continued at a healthy pace even after the end of the miracolo economico, and favored the poor. By any measure – top income shares, the Gini index, or the downward slope of the GIC – inequality diminished.
In the quarter century since, this trend has been entirely reversed. The real incomes of the poorest 30% have decreased in absolute terms; those of the next 60% (the middle class) have stagnated; and only the richest 5-10% have seen significant growth.

The findings in Figures 3 and 4 are consolidated, robust results. We turn now to more speculative comparative exercises that illustrate both the promise and some perils of our approach. We first turn the time travelling welfare analyst’s dial all the way back to 1427, the year of a renowned tax survey (catasto) in Florence. The returns for all households in the city survived, and constitute a precious source of data on family wealth, size, and occupations. First subjected to computer analysis by Herlihy and Klapisch-Zuber (1985), the catasto data have been revisited more recently by

Source: Amendola and Vecchi (2016).
Milanovic, Lindert, and Williamson (2011). Property income is directly available from the *catasto*; the authors estimate labour income from the head of household’s occupation and scattered observations on wage rates, combined with assumptions about working days per year, the contributions of other family members, and the value of owner-occupied homes. Figure 4 illustrates the resulting distribution of family income per capita, alongside the 1861 Italian distribution, reprised from Figure 2.

**Figure 4. Four centuries and a half of decline**

![Cumulative Distribution Functions](image)

Source: authors’ elaboration on datasets kindly made available by Peter Lindert (Florence 1427) and Amendola and Vecchi (Italy 1861).

The retrogression between 1427 and 1861 is dramatic; can this possibly be right? It *does* correspond to the historiography of Italian economic decline, and to what we know about real wages, urbanisation, and nutrition. Malanima’s (2011) estimates of North-Italian GDP per capita in the long run indicate a 22% decline between 1427 and 1861. But the contrast is exaggerated, for we are comparing the richest city of Renaissance Italy’s richest region with all of Italy in 1861 including the poorest agricultural laborers.
of its poorest villages. The income estimates are also subject to wide margins of error. Italian incomes of 1861 are directly observed, but for only a sample, while Florentine incomes are only estimates, though available for every household in the city. The conversion of 1427 florins to 2010 euros brings with it a further set of challenges.

As interesting, and as difficult, as comparisons over time are those across countries. Figure 5 illustrates a comparison of Italy and the UK in the first decade of the twentieth century. The British data, already presented in Figure 1, are a large collection including more than 900 households from the Board of Trade’s 1904 study of urban working men’s families (Gazeley and Newell, 2011), the families of some 200 railway clerks representing the lower middle class (Scott and Walker, 2015), and 920 family budgets gleaned from 19 different studies covering a range of locations, sectors of employment, and locations (A’Hearn, Di Battista, and Vecchi, 2016). The three-sector, two-region post-stratification weighting is based on an interpolation of the 1901 and 1911 population censuses.

We express household incomes in 1905 domestic currency terms, using domestic price indices, but then face the difficult task of picking a PPP exchange rate to convert Italian lire to British pence. Williamson (1995) estimated this at 25.9 lire/pound, quite close to the market exchange rate of 25.2 that is used in Figure 5. But Williamson’s calculation can almost certainly be improved, for it is based almost entirely on 13 food items common to the seven countries he compares. First, we should consider that dietary patterns differed widely across these countries. Bread, for example, was 30% of Italian food expenditure, just 14% of British, and 19% of the seven-country average used to compute PPP standards. Or take butter: 13% of British, 2% of Italian, and 11% of average expenditure. The only non-food item in the calculation is rent, which unfortunately was not available for Italy and had to be imputed.

---

22 Lindert et al. estimate per capita income in the city at 153 Florentine lire. Malanima’s (2011) estimate for per capita GDP in all of North Italy is 63. This gives a sense of the relative wealth of the city of Florence.

23 For comparability we convert the 1427 figures from florins to euros of 2010 using Malanima’s (2011) North-Italian price index, additionally converting 1427 florins to 1427 Florentine lire at the rate of 4.15:1, Florentine lire to Italian lire at the 1861 rate of 3.8/4.5, and Italian lire to euros at the rate 1/1936.
Data on rents is available for Italy, but only for the cities, where accommodation was expensive. An important item in British budgets (14% relative to food expenditure according to a 1918 study) that is omitted from the calculations is fuel; cheap coal was used in large quantities. In Italy fuel was considerably less important in the budget, but consisted of expensive firewood. Inclusion of fuel (making it part of the set $G$, using the notation of Section 5) with an expenditure weight ($q$) equal to a British-Italian average thus makes Italian prices look expensive. In preliminary calculations, A’Hearn et al. (2016) find that the combined effect of all these changes (two-country expenditure weights, Italian urban rents, and inclusion of fuel) result in the lira looking rather overvalued at 25.2 per pound; PPP rates come out closer to 40.

The most striking feature of Figure 5 is the crossing of the two curves. The 80th and 90th percentiles of the Italian distribution are at income levels clearly above their British counterparts. In part this may result from overvaluing the lira in our currency conversion. But it is certainly affected by the non-random nature of the underlying British budget data. Missing from this otherwise-impressive collection are the upper middle class and the wealthy. Inclusion of such families would shift the British
distribution to the right. Still, we can make reliable comparisons of the lower tails. Below about 10p per person per day, and particularly in the range of a plausible poverty line, which would lie somewhat lower (in the shaded gray region), there are many more Italian than British families. Observing the vertical distance between the distributions we see that at least 10% more Italian families have daily per capita incomes of 5p or less, relative to British. A higher rate of lire per pound or better coverage of well-off families would shift the British distribution down or to the right, and could only exacerbate this disparity.

The examples presented illustrate both the power and some of the perils of historical household budget analysis. It has the potential to generate subtle insights into the changing distribution of welfare across households, on a comparable basis, over long periods of time, and across countries. At the same time it requires a clear-eyed assessment of the limitations of the sources, depends to a considerable extent on non-sample information such as price indices, and must be checked for corroboration against alternative estimates.

8 Call to arms

We have argued three points about historical household budgets in these pages. First, they represent the ideal type of data for addressing both micro questions about families and macro questions about distribution. Second, they are far more numerous and widespread than generally supposed. Third, the modern welfare analyst travelling back in time already has a toolkit well-suited to dealing with the problems thrown up by historical family data. All this leads us to conclude that historical household budgets, far from a subject for antiquarians, have a bright future.

The challenges of this research agenda should not be underestimated. The technical difficulties to which we have devoted most of our attention here – non-representativeness, non-standardization, and currency conversions over long time periods or across countries – are capable of solution, but only at the cost of painstaking work with the sources, engagement with the relevant statistical techniques, and the search for complementary information on population structure and the prices of consumption items. But if the challenges are great, so too is the payoff. Should the HHB
Project’s aim of making the micro data and the welfare indicators derived from them comparable over time and across countries be realised, we will be in a vastly better position to understand the historical development of living standards, the distribution of welfare, and how they responded to technical, institutional, and policy changes.

So we conclude with a call to arms. There is a need for every sort of contribution, from the discovery of apparently insignificant budget sources to the perfection of minor variations on a technique for estimating distributions to the development of new historical narratives and interpretive frameworks. Lend your talents and energies to the fight.
References


Atkinson, A. B., and T. Piketty (eds.), (2007), Top incomes over the twentieth century: a contrast between continental european and english-speaking countries. OUP
Oxford.


Burkhauser, R.V., N. Hérault, S.P. Jenkins, and R. Wilkins (2016), ‘What has been happening to UK income inequality since the mid-1990s? Answers from reconciled and combined household survey and tax return data’, NBER Working paper 21991.


Suomalaisen Kirijallisuuden Seuran Toimituksia.


Moehling, C. M. (2005), ‘She Has Suddenly Become Powerful: Youth Employment and Household Decision Making in the Early Twentieth Century’, *The Journal of


Scott, P., and J.T. Walker (2015), 'Demonstrating distinction at ‘the lowest edge of the black-coated class’: The family expenditures of Edwardian railway clerks’, *Business History*, 57(4) 564-588.


Stigler, G. J. (1954), ‘The early history of empirical studies of consumer behavior’. *The


