

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Explorations in Economic History

journal homepage: www.elsevier.com/locate/eehIncome inequality in Eastern Europe: Bulgaria and Czechoslovakia in the twentieth century[☆]Stefan Nikolić^{a,*}, Filip Novokmet^{b,c}, Piotr Paweł Larysz^d^a Loughborough University, Epinal Way, LE11 3TU, Loughborough, United Kingdom^b University of Zagreb, Faculty of Economics and Business, Trg J.F. Kennedy 6, 10000, Zagreb, Croatia^c World Inequality Lab, France^d Free University of Berlin, School of Business and Economics, Boltzmannstr. 20, 14195 Berlin, Germany

ARTICLE INFO

JEL classification:

D31
E24
E25
J31
N33
N34

Keywords:

Income inequality
Social tables
Bulgaria
Czechoslovakia

ABSTRACT

This article provides novel estimates of long-term income inequality in Bulgaria and Czech Lands/Czechoslovakia in the twentieth century. Relying on newly-constructed datasets and the social tables approach, we measure inequality between salient social strata. We find that Czechoslovakia was significantly more unequal than Bulgaria before 1945. Inequality converged to similarly low levels under socialism. Decomposition analysis by social classes reveals that different levels of inequality in the first half of the century were principally driven by higher within social-class inequality in Czechoslovakia, owing to a more stratified industrial society; whereas a low dispersion within the dominant agricultural sector held down the within social-class component in Bulgaria. A dramatic fall in total inequality after 1945 was a result of the social revolution that encompassed the virtual disappearance of between social-class inequality and a marked reduction in within social-class inequality. Our findings point to the critical role of institutional and political factors in driving inequality in Eastern Europe throughout the twentieth century.

1. Introduction

The historical evolution of income inequality in Eastern Europe is scantily documented.¹ This is unfortunate given the dramatic socio-economic and political changes that singularly marked the twentieth century in the region, and could serve as a historical laboratory to investigate central developmental questions: How does inequality evolve during economic development and which structural forces shape it? How does development impact social structure? Which social groups benefit or lose from modern economic growth?

[☆] We thank the participants of the *Explorations in Economic History Workshop on Wealth and Income Inequality Around the World* (October 6-7, 2023, Federal Reserve Bank of Chicago), Steven Nafziger, the discussant of our paper, the editors and referees for their invaluable feedback. We are grateful to Martin Ivanov, Branko Milanović and Thomas Piketty for their helpful comments. We thank María Gómez León and Giacomo Gabbuti for their advice on social tables. We are grateful to Tri Hoang for his generous help with coding. Nikolić and Larysz acknowledge support from the European Research Council Horizon 2020 Starting Grant no. 803644. "Spoils of War: The Economic Consequences of the Great War in Central Europe".

* Corresponding author.

E-mail addresses: s.nikolic@lboro.ac.uk (S. Nikolić), fnovokmet@efzg.hr (F. Novokmet), piotr.larysz@fu-berlin.de (P.P. Larysz).

¹ Promising work on historical inequality in Eastern Europe has been done, but available series either focus on the development at the top of the income distribution (Novokmet, 2017, 2023; Novokmet et al., 2018; Bukowski and Novokmet, 2021), are confined to a specific point in time (Lindert and Nafziger, 2014; Vinski, 1967), or start relatively late (e.g. in the 1960s (Atkinson and Micklewright, 1992)). Studies of the long-run aiming at global coverage approximate inequality in the whole of Eastern Europe based on sporadic estimates available for certain countries and years (Bourguignon and Morrisson, 2002; Van Zanden et al., 2014).

<https://doi.org/10.1016/j.eeh.2024.101594>

Received 31 January 2024; Received in revised form 14 May 2024; Accepted 21 May 2024

Available online 31 May 2024

0014-4983/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

This paper aims to address these important questions by constructing and analyzing a comprehensive dataset on long-run income inequality in Bulgaria and Czechoslovakia during the twentieth century. Our country choice is motivated by comparing inequality levels and trends in the most and the least developed parts of Eastern Europe, which makes them particularly suitable case studies from the comparative development perspective.² We track inequality during a tumultuous period that included the two world wars, state formation, the Great Depression, German occupation of Europe, and Communism.

We highlight two main findings on long-run inequality. First, income inequality was significantly higher in Czechoslovakia than in Bulgaria in the first half of the twentieth century (the Gini coefficient was between 40–45 in Czechoslovakia and around 30 in Bulgaria). Second, income inequality sharply declined in both countries with the introduction of the communist system after the Second World War, converging to record low levels under socialism (the Gini coefficient reached levels slightly below 20). In addition, the Great Depression and the Second World War stand out as significant short-term shocks to income inequality. During the Great Depression, inequality increased in Bulgaria and fell in Czechoslovakia, owing to structural differences of the two economies. Towards the end of the Second World War, inequality decreased in German occupied Czechoslovakia.

We estimate income inequality using the social tables approach. Social tables are data tabulations enumerating salient socio-economic groups with their average incomes (Milanovic et al., 2011, p. 256). Social tables have a long tradition in economic history. In modern literature, pioneering social tables were created by Lindert and Williamson (1982), and to date no less than 41 social tables have been produced (Milanovic, 2018).³ While most social tables are constructed for benchmark years, following Rodriguez Weber (2015), scholars have recently started to construct annual social tables. Gómez León and De Jong (2019) estimated income inequality in Britain and Germany from 1900 to 1950 based on social tables, and Gómez León and Gabbuti (2024) have done the same for Italy. Social tables are the main tool for estimating inequality in periods for which tax data or household surveys are lacking. Moreover, social tables aim to draw a full picture of a society's income distribution.

Social tables enable us to examine how changes in the size and incomes of social groups affected inequality. We are thus able to consistently track, for the first time, the secular developments in social stratification in the studied countries, as well as to ascertain how social changes impacted long-run inequality trajectories. Structural change, compression or expansion of worker's wages, a growing public sector, collectivization of agriculture and nationalization of industry mattered in both studied countries at certain periods.

To construct annual social tables, we muster a comprehensive new dataset on employment structure and incomes. For both studied countries we start on the eve of the First World War, cover the interwar period, the Second World War, and most of the socialist period. In total, we cover 53 years in Bulgaria, and 57 years in Czech Lands/Czechoslovakia. Based on population censuses, we track the labor force by consistent social groups over time. Workers, employees, and self-employed are joined by members of cooperatives under socialism.⁴ We distinguish social groups within more than twenty economic branches spanning all three major economic sectors. We draw on a multitude of historical sources for average incomes of salient social groups. Sources include statistical yearbooks, labor statistics, health and accident insurance reports, tax statistics, household budget surveys, microcensuses, and national income studies.

Compared to the existing literature, we go a step further to capture inequality within social groups. With few social groups and economic sectors, unobserved income differences within social groups can be potentially huge. We apply the most granular social and economic differentiation possible by assigning incomes across many industries to multiple social groups. This way we minimize income differences within social groups. Further, we innovatively use additional data for social groups with potentially large within-group income differences. For example, we use land surveys to estimate the income distribution of farmers. We draw on personal income tax data for income differences among self-employed. We apply income distributions of employees from employer surveys. And we distribute members of cooperatives by income based on household budget surveys. Lastly, for social groups where only average incomes are available, we employ the log-normal distribution to estimate within-group inequality.

We apply the social class analysis to explain both the documented inequality levels and trends. We formally assess the role of social structures for inequality by conducting a standard inequality decomposition into constituent within and between inequality. We find that the different inequality levels in the two studied countries in the first half of the twentieth century were the result of a more “developed” social structure in Czechoslovakia. This entailed a relatively larger contribution of within inequality in Czechoslovakia due to the more sizeable shares of modern (industrial) classes that were also characterized by higher within inequality. On the other hand, the importance of the between component – reflecting differences in the mean income between classes – was roughly similar in the two countries. Concretely, in pre-industrial Bulgaria, the relatively egalitarian distribution among the most populous group of independent farmers attenuated the sizeable income distances between groups (i.e. within inequality was constrained).⁵ In Czechoslovakia, different social classes were more sizeable (more evenly represented in the population) and at the same time all exhibited relatively higher within inequality, hence the contribution of within component was more pronounced.⁶ The

² Around 1930, Czechoslovakia had the highest and Bulgaria the lowest share of industrial employment in Eastern Europe (ca. 40 and 8% respectively, cf. Fig. 2).

³ For example, social tables have recently been constructed for the USA (for 1774, 1850, 1860, and 1870) by Lindert and Williamson (2013, 2016), and revised for England (for 1688, 1759, 1798, 1846, and 1867) by Allen (2019).

⁴ Population censuses clearly distinguish various social categories. We group certain social categories to ensure consistency across censuses in the long run.

⁵ However, it should be emphasized that land inequality was not necessarily low in pre-industrial societies, and correspondingly the contribution of within inequality to overall inequality was not relatively egalitarian. We discuss the critical role of land inequality for pre-industrial inequality.

⁶ For example, the inequality among workers, as the largest social class in pre-communist Czechoslovakia, was high due to pronounced gender wage inequality or marked wage dispersion between industrial branches; inequality among independents outside agriculture was markedly right-skewed due to a large income gap between capitalists and the rest of self-employed; inequality among independent in agriculture was high due to very high land concentration, etc.

social revolution that occurred during communism had a first-order impact on income inequality. The total income inequality came to be almost exclusively determined by the within-inequality component, itself reduced to low levels (especially in Czechoslovakia). Inequality in mean income between social groups practically vanished, partly in an attempt to reach a classless society.

We confront our results with existing theoretical frameworks of the inequality evolution during the development process, in particular in relation to mechanisms underlying the influential Kuznets' hypothesis – that of sectoral reallocation of labor and cumulative effects of concentration of saving. For this purpose, we extend the analysis by comparing patterns of top income shares in Bulgaria and Czechoslovakia.⁷ We explain observed inconsistencies with the predicted dynamics of the paradigmatic model in its overly deterministic emphasis on economics forces as determinants of inequality. We suggest instead the critical role of institutional and political factors in shaping long-run income inequality in Eastern Europe during the twentieth century.

The next section discusses the historical and conceptual background. Section 3 explains the social tables approach to estimating income inequality. Section 4 describes our data on socio-economic structure and incomes. Section 5 presents estimates of income inequality in Bulgaria and Czechoslovakia in the twentieth century. Section 6 compares our results with the international literature. The last section offers our conclusions.

2. Historical and conceptual background

The economic development in Eastern Europe in the “long” twentieth century was turbulent (see Fig. 1). The two countries studied, Bulgaria and Czechoslovakia, both passed through critical historical junctures of the two world wars, the rise and fall of communism, globalizations and disintegrations, among others, which all entailed dramatic socioeconomic and political transformations. At the beginning of the studied period, at the turn of the twentieth century, the two countries belonged to markedly different political, socioeconomic and cultural spheres, being for centuries a part of large multinational empires—the Czech Lands developing in the Habsburg Empire (since the sixteenth century), Bulgaria in the Ottoman Empire (since the late fourteenth century). Countries' economies accordingly assumed distinctive features, for example, visible in different institutions, administrative capacities, land tenure arrangements, the nature of the proto-industrialization, etc.

By the First World War, the Czech Lands underwent a critical economic transformation and largely industrialized in a “natural” (capitalistic) way. In these lands commercial agriculture flourished and Habsburg industry thrived, accounting for almost two-thirds of the monarchy's industrial capacity (Klein et al., 2017). Moreover, the crown lands boasted one of Europe's most educated workforces (Cvrček, 2020). Bulgaria was, in contrast, among the poorest countries in Europe. The prevailing historical narrative describes the country's experience in the first half of the twentieth century as that of failed economic modernization (Gerschenkron, 1962; Lampe and Jackson, 1982; Lampe, 1986; Palairat, 1997).⁸ Subsistence agriculture in tandem with the gravest rural overpopulation, and the small and sluggish industry, kept Bulgaria in the low-growth equilibrium.

The aftermath of the First World War brought the dissolution of the Habsburg Empire and the establishment of the First Czechoslovak Republic. Bulgaria, de facto independent from the Ottoman Empire from 1878, came out of the Great War (and previous Balkan Wars) burdened with war reparations and high sovereign debt. Both countries grappled with a difficult political situation coupled with ethnic tensions. Czechoslovakia further faced the challenge of nation building. The Great Depression hit these Eastern European countries hard compared to the rest of Europe (Marcus et al., 2021), and their overall growth experience in the interwar period was relatively poor.

The aftermath of the Second World War signified communist accession to power in both countries and falling into the Soviet influence zone (“Iron Curtain”). This entailed a full-scale transformation of the economy and a rigid adoption of the Soviet economic model. Both countries came to be known as bastions of the orthodox communist model (ensured with direct Soviet military intervention in Czechoslovakia in 1968). This primarily meant the implementation of central planning, nationalization of industries, and collectivization of agriculture.

Inequality and development. The structural change from agriculture to industry (i.e. industrialization) may be taken as the most robust indicator of initiated economic development (Kuznets, 1966, 1973). A fundamental asymmetry between sectors lies at the heart of the famous dual-economy model (Nurkse, 1953; Lewis, 1954; Ranis and Fei, 1961). The model builds on a notion that high-income countries underwent a structural change, with the massive reallocation of labor (and other resources) from the traditional, low-productivity agricultural sector to modern (urban and industrial), high-productivity sectors. Most importantly, differential sectoral productivities postulate aggregate growth-enhancing effects of sectoral labor reallocation.⁹

The stylized dual-economy model provided an operating framework for Kuznets' model of inequality evolution during the development path. Kuznets (1955) famously hypothesized an inverse-U evolution of inequality, according to which inequality first increases and then declines as the country develops. He proposed a mechanism of structural reallocation of the population from the traditional agricultural sector to the more productive modern sectors that drives the upward swing of the inverse-U trajectory.

⁷ Our other data contribution is to construct the first historical series of top income shares in Bulgaria for the 1921–1946 period.

⁸ “Revisionist” views attenuate these harsh assessments stating that important preconditions were made before the Second World War for the eventual communist industrialization of Bulgaria (Ivanov and Tooze, 2007; Ivanov and Kopsidis, 2023). Recent research on neighboring Serbia found that both real wages (Milanović and Mijatović, 2021) and GDP per capita (Mijatović and Zavadžil, 2023) were stagnant between 1860 and 1910, thus supporting the traditional underdevelopment thesis.

⁹ It has been often assumed that developing countries are featured by surplus labor in agriculture (i.e. marginal product of labor is close to zero), which attracts workers to modern, technologically superior sectors (without a loss of agricultural output). This inflow of workers keeps wages in the industry low, allowing in turn rapid capital accumulation there (which contributes further to rising labor productivity).

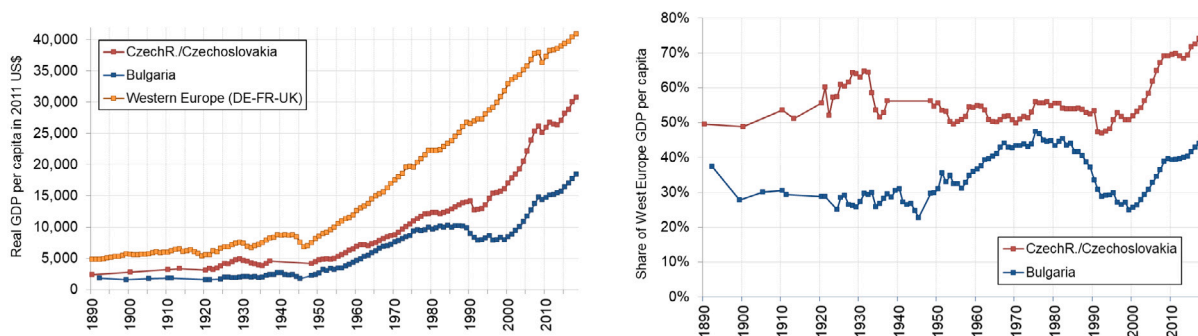


Fig. 1. Long-term GDP per capita in Czechoslovakia and Bulgaria, 1890–2018.

Notes: Up to 1990 Czech Lands/Czechoslovakia; since 1990 the Czech Republic. Western Europe is the unweighted average of France, Germany and the United Kingdom.

Source: Own construction from Bolt and Van Zanden (2020).

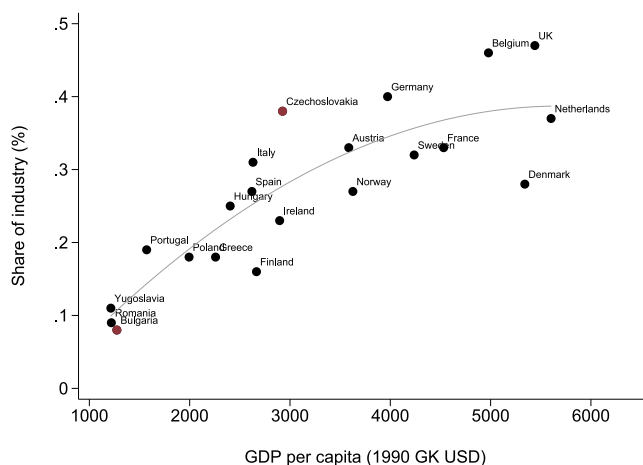


Fig. 2. Share of employment in industry and GDP per capita in Europe in the interwar period.

Notes: Shares of employment in industry are for 1930; GDP per capita for 1930 (in 1990 international Geary-Khamis dollars).

Source: Share of employment in industry: Buyst and Franaszek (2010, Table 9.1, p.210). GDP per capita: Bolt and Van Zanden (2020).

Kuznets postulated different inequality levels in the two sectors (thus extending the paradigmatic dual-economy model, e.g. as conceived by Lewis (1954)), with lower inequality in agriculture and higher in industry – hence, a reallocation of the population during industrialization leads both to an increase in the aggregate output and of inequality (Anand and Kanbur, 1993). In the later stages of development, inequality falls once the surplus labor is fully absorbed and wages increase in the rural sector (Ravallion, 2016), as well as inequality declines within the modern sector (see generally Milanovic (2023)).

The comparative experience of Bulgaria and the Czech Lands provides a particularly useful setting to investigate the validity of Kuznets' propositions. Both the timing and nature of industrialization were markedly different in the two countries. Different development levels during the study period imply that each country may be seen at different positions on the hypothesized Kuznets' curve. Recorded inequality patterns may be thus indicative of the assumed evolution and mechanisms operating during different phases of development – i.e., Bulgaria for the growth take-off and the rising part of the curve; the Czech Lands for the more mature stages of development presumably accompanied by the decline in inequality.

At the turn of the twentieth century, when our analysis starts, we can look at them as being representative of a European core and a periphery (Aldcroft, 2016). Fig. 2 shows that Bulgaria was among the poorest and the least industrialized European countries during the interwar period. In the first half of the twentieth century, Bulgaria (and South-Eastern Europe in general) could be seen as a prototypical development country, which did not still commence its economic “take-off”.¹⁰ After their accession to power,

¹⁰ Lampe and Jackson (1982) thus point out that South-Eastern European “periphery” represented the original development countries. In fact, the earliest contributions to the dual-economy model were to a considerable extent inspired by the development predicaments of South-Eastern Europe (Rosenstein-Rodan, 1943). Surplus labor was perceived as characteristic of South-Eastern Europe in the first half of the twentieth century due to rural overpopulation (see footnote 9).

communists forced industrialization and structural change following the Soviet example of the 1930s, which, in its essence, was a variant of the dual-economy model.¹¹

We examine the inequality dynamics during Bulgaria's industrialization, but find no evidence that the take-off was accompanied by rising inequality. As we show, structural transformation was actually accompanied by a dramatic decline in inequality. Reasons for this inconsistency with the predicted upward swing of inequality in the paradigmatic model lie in its overly deterministic interpretation of forces driving inequality, in particular, its sole focus on the impact of a sectoral reallocation of workers on labor market outcomes. More generally, the Kuznets process neglects the critical role that institutional and political factors play in shaping inequality in the long run, both of labor income and other income sources (Lindert, 2000; Piketty, 2006). Their decisive role for inequality during Bulgaria's take-off is clearly manifested in the communist (almost complete) bureaucratic control of employment and wage setting, which allowed them considerable discretion in determining wage differentials. Similarly, "institutional" investment in the massive expansion of education increased the supply of skilled workers (that was plausibly rising in tandem or faster than the industrialization-induced demand for skill).¹²

In contrast, Czechoslovakia – and the Czech Lands in particular – had already achieved relatively high living standards by the interwar period (e.g. similar to Austria, Italy, or Norway) and undergone considerable structural change with one of the highest levels of industrial employment in Europe (Fig. 2). Accordingly, we can analyze inequality patterns in the industrialized Czech Lands in a setting of the conjectured peak (or the turning point) and the downward part of the inverse-U curve.¹³

We document a dramatic secular decline of inequality in Czechoslovakia in the post-Second World War decades, however, as in the case of Bulgaria's take-off, it had little to do with the dual-economy processes, as postulated by Kuznets. Wage inequality declined primarily because of comprehensive institutional reforms of the labor market and pronouncedly egalitarian wage policies. More fundamentally, income inequality declined due to the dramatic deconcentration of capital income as communist nationalizations and expropriations completely eliminated private ownership of wealth (Novokmet, 2023). As first shown by Piketty (2001, 2003), a decline in inequality in Western countries was largely a result of the major shocks to high capital incomes (Atkinson and Piketty, 2007, 2010; Atkinson et al., 2011), rather than caused by the Kuznets' mechanism. In this respect, the Czechoslovak downward trajectory may be simply seen as the extreme version of inequality development in Western developed countries.

Finally, it should be mentioned that Kuznets (1955) identified two major sources of inequality during the development process. In addition to the process of sectoral reallocation of population discussed thus far, he also mentions the concentration of saving among high-income groups. Given that top incomes have a higher marginal propensity to save, there is a tendency towards higher wealth concentration, which leads in turn to higher capital income concentration. Correspondingly, an ongoing development should lead to a rising concentration of capital income and rising top income shares.¹⁴ We investigate this mechanism by comparing top income shares in Bulgaria and Czechoslovakia before communism. In line with the presented logic, we find indeed notably higher top income shares in industrialized Czechoslovakia, which were largely composed of capital income.

3. Methods

Social tables establish the number of individuals in salient social classes and estimate their average incomes. Social tables are a powerful tool for analyzing historical inequality before the widespread introduction of individual taxation or household surveying. Indeed, before the 1950s, there were relatively few household surveys, and the ones existing for the countries we study cannot be taken as representative of the total population. At the same time, tax records focused on top income earners, leaving out a considerable portion of population at lower income levels. Social tables, by contrast, aim to capture a more comprehensive picture of a country's socio-economic structure.

We follow the approach of Milanovic et al. (2011) and other scholars, to construct our social tables for the gainfully occupied population. Thus, we measure inequality of individual incomes per economically active person.¹⁵ This includes between 50 and 60 percent of the total population earning a market income. We draw on occupational censuses to categorize active population by social groups, economic sectors, and gender. Building social tables for the twentieth century provides for a relatively large number of observations compared to pre-modern societies. This is both because of the higher economic specialization and social stratification

¹¹ The industrialization process was forced in the sense that independent peasants were expropriated and quickly turned into wage earners, either compelled to join the collective farms or absorbed by expanding state-owned industry.

¹² Importantly, the predominantly state ownership of the means of production entailed that the growing importance of capital income during industrialization appertained fully to the state, thus removing this inegalitarian pressure from interpersonal inequality.

¹³ Our series starts at the beginning of the twentieth century, during a period that historical research has often associated with the peak of the Kuznets' curve, notably in the development context of Central Europe (especially of Germany (Dumke, 1991; Grant, 2005; Tilly and Kopsidis, 2020)). The Czech Lands are frequently seen as developing in tandem with neighboring German and Austrian Lands, and their economies consequently assuming broader common features (see also Fig. 2).

¹⁴ The more so as development in a dual-economy model is supposed to be accompanied by the rising share of capital income in national income (Ravallion, 2016).

¹⁵ It is hence the data availability that dictates the choice of our benchmark unit of observation as (economically active) individual adult. We correspondingly analyze the distribution of individual incomes for social categories of workers and employees (i.e., we assume that there is no income sharing within households of workers or employees). However, several adjustments were required due to data constraints. Income data for independents in agriculture before the Second World War is estimated at the household level from land censuses, and we needed to split income between adults within a household. We have applied an 'equal-split adult' unit for independents in agriculture, according to which household income is equally distributed within households (Alvaredo et al., 2016). We assume an equivalent approach for cooperative farmers and independent farmers during socialism, whose income is obtained from Household Budget Surveys (HBS); i.e., we equally distribute household income from HBS to adult individuals.

in modern societies, and the more detailed recording of data in official statistical sources. For example, each major economic sector consists of multiple branches within which we are able to distinguish several social groups, such as workers, employees, or entrepreneurs. Social tables, thus, truly provide a most detailed look into the socio-economic fabric of a country and inequality of income thereof.

Social tables, however, have their limitations. The most pertinent limitation concerns inequality within social groups. Although social tables enumerate individuals in each social group, individual incomes remain unobserved. Instead, average incomes are imputed to each social group. This may downward bias inequality if income variation within a social group is high. We adopted several strategies to mitigate this issue.

First, based on population censuses, we apply the largest possible level of social and economic disaggregation for a given historical context. In the first half of the twentieth century, censuses would typically differentiate between “blue-collar” workers, “white-collar” employees, and the self-employed (called “independents” by the census), who depending on the economic sector, were farmers, industrialists, business owners, free professionals, or other entrepreneurs.¹⁶ During socialism, with the nationalization of industry and collectivization of agriculture, independents all but disappeared, while cooperative members emerged as a new social category.¹⁷ Importantly, we not only account for differences in incomes between social groups, but also across economic sectors. For example, workers in agriculture or traditional industries such as textiles would typically earn considerably less than those in modern industries (chemicals, for example). Detailed, branch-level statistics allow us to account for these income differences.

Next, we innovatively combine land distributions and tax records to estimate incomes within social groups where within-variation is expected to be large.¹⁸ In Bulgaria, for example, farmers were the largest social group. Assigning the same average income to all farmers would mask potentially large income differences between farmers with land plots of different size. We take into account how land was distributed among farmers to capture inequality within this social group. Another example are entrepreneurs in industrial Czechoslovakia. Average incomes would hide differences in earnings between small business owners and large industrialists. To remedy this issue, we employ tax data to separate entrepreneurs in each sector into income classes. This way we account for inequality of income among the top income earners.

Lastly, we do not generally have information on the within-group income dispersion of workers (and of employees before 1945). [Modalsli \(2015\)](#) offers an important discussion of the issue specifically related to the social table method. He applies the log-normal distribution to the existing estimates of social tables and obtains significantly higher estimates of inequality as compared to those before the correction. We similarly deal with this issue by assuming that incomes within a particular social group (for which we only know mean income) follow the log-normal distribution. This is a standard assumption given the well-known empirical regularity that the size distribution of income is well described by the log-normal distribution (specifically, since it is positive and right-skewed). Hence its frequent application to model the income distribution.¹⁹

Parametrizing the log-normal form to our data is straightforward as this distribution is described by two parameters, the mean (μ) and the standard deviation (σ) of the log income. As we know the mean income of each social group, the standard deviation is assumed based on a conventional value for the coefficient of variation (CV); i.e. standard deviation equals the mean times the coefficient of variation. We opt for the relatively lower values of the coefficient of variation, of 0.5 (for example, relative to values used by [Modalsli \(2015\)](#)), since we model the within-variation for smaller groups, that is, we have much more detailed distinction of occupational groups than is usually the case in the literature.²⁰ Finally, from thus simulated log-normal distribution for a given occupational group (e.g. for workers in metallurgy), we create five composite observations based on the number of standard deviations from the mean ([Fig. 3](#)).²¹

4. Data

To construct social tables and estimate long-term income inequality, we muster a comprehensive new dataset on the socio-economic structure and incomes in Bulgaria and Czech Lands/Czechoslovakia spanning the period from 1911 to 1980. In this section, we describe the main sources and illustrate key data features, while the full list of sources and details on data building are documented in [Appendix](#).

[Table 1](#) summarizes our dataset. In both countries we consistently track five social groups. As explained in [Section 3](#), we divide each social group into a number of sub-groups to capture within-group variation. Social groups are further distinguished by more

¹⁶ Censuses clearly distinguished workers from employees. Workers included laborers, apprentices, and others who performed manual work in agriculture or industry, or entry-level jobs in services. Employees included officials, clerks, and others tasked with administrative and management jobs spanning all major economic sectors.

¹⁷ To be clear, the population census first introduced cooperative members as a social category in 1956 in Bulgaria and in 1961 in Czechoslovakia.

¹⁸ It should be noted that [Lindert and Nafziger \(2014\)](#) used tax and land inequality data to generate their estimate of Russian inequality. Our innovation is in using these data in a new way.

¹⁹ Log-normality of income is explained in relation to the Gibrat's law of proportionate growth, according to which income of a given group is determined by series of random multiplicative shocks giving rise to a log-normal form ([Gibrat, 1931](#); [Kalecki, 1945](#); [Aitchison and Brown, 1957](#)).

²⁰ The log-normal form described reasonably well the income distribution in former socialist countries ([Lydall, 1968](#)). However, since the dispersion of income in socialist countries was found to be relatively lower, we have modeled the log-normal form with a CV of 0.25 to simulate the within-group inequality during the socialist period.

²¹ From the lowest to highest income observation: (i) less than 1σ below the mean [$0, \mu - 1\sigma$]; (ii) from 1σ below the mean to 0.5σ below the mean [$\mu - 1\sigma, \mu - 0.5\sigma$]; (iii) from 0.5σ below the mean to the mean [$\mu - 0.5\sigma, \mu + 0.5$]; (iv) from the mean to 1σ above the mean [$\mu, \mu + 1\sigma$]; (v) more than 1σ above the mean [$\mu + 1\sigma, +\infty$]. We call these composite observations (rather than quintiles) since they are not necessarily of the same size.

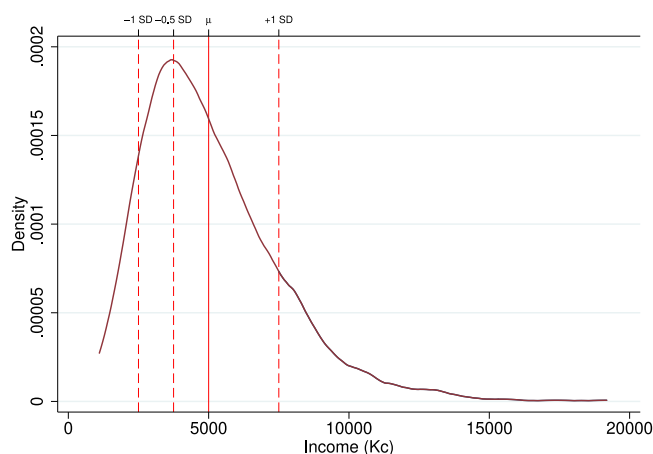


Fig. 3. Simulated lognormal distribution with $CV=0.5$ ($\mu=5000$ and $\sigma=2500$).
Source: Own illustration.

Table 1
Dataset summary.

Social groups	Within groups (no. of groups)	Sectors (no. of branches)	Period (no. of years)	N
Bulgaria				
Workers	log-normal (5)			10,600
Employees	employer census (5)	Agriculture,	1911, 1921–1946,	10,600
Independents in agriculture	land size (8)	Industry (15),	1955–1980 (53)	636
Independents outside agriculture ^a	log-normal (5)	Services (4)		5,130
Members of cooperatives ^b	HBS (5)			4,940
				31,906
Czech Lands/Czechoslovakia				
Workers	log-normal (5)			13,110
Employees	microcensus (5)	Agriculture,	1913, 1921–1945,	13,110
Independents in agriculture	land size (6)	Industry (15),	1950–1980 (57)	684
Independents outside agriculture ^a	tax income (7)	Services (7)		8,008
Members of cooperatives ^b	HBS (5)			6,820
				41,732

^a Independents outside agriculture disappear after 1945.

^b Members of cooperatives appear after 1945.

Notes: N is the number of observations calculated as the sum product of each social group and the corresponding number of within groups, economic branches, gender, and years. Before 1945, we estimate within inequality of employees applying the log-normal. After the Second World War, we rely on results of Bulgarian employer censuses (reported in Statistical Yearbooks) and the 1955 and 1970 Czechoslovak microcensus. HBS stands for household budget surveys.

than twenty economic branches spanning all three major sectors, and by gender. Our data cover a benchmark year before the First World War, most of the interwar, the Second World War, and the socialist period. In total, we have 31,906 observations for Bulgaria and 41,732 observations for Czech Lands/Czechoslovakia.

4.1. Socio-economic structure

Population censuses reveal a country's socio-economic structure. We collected labor force data from population censuses of Bulgaria in 1910, 1920, 1926, 1934, 1946, and 1956, as well as censuses of Czech Lands in 1910, and Czechoslovakia in 1921, 1930, 1950, and 1961. While earlier censuses are available, it proved difficult to find all the necessary accompanying incomes to reconstruct social tables for the preceding period. We, therefore, settled on constructing social tables starting from the eve of the First World War.

We consistently track socio-economic groups throughout most of the twentieth century. This allows us to estimate inequality before and after the First World War, during the tumultuous interwar, and uniquely both for Bulgaria and for the Protectorate of Bohemia and Moravia (that is the Czech Lands under German occupation) during the Second World War. Furthermore, our analysis covers the socialist period. To the best of our knowledge, we construct the first social tables for socialist countries. We can, thus, evaluate how inequality changed in the transition from a capitalist to a communist mode of production. For the socialist period, the first socialist census serves as the benchmark from which we reconstruct annual labor force movements with data reported in statistical yearbooks.

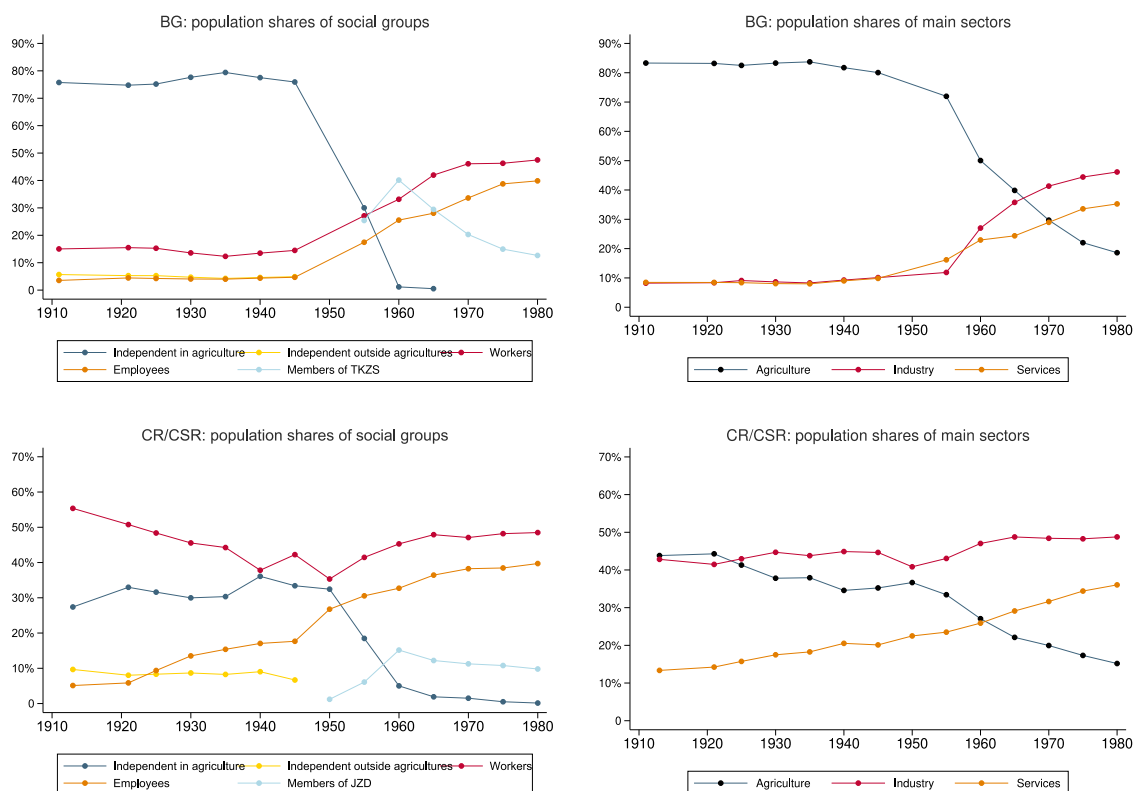


Fig. 4. Labor force shares by social groups and economic sectors.

Notes: Figures illustrate labor force shares by social groups (left) and economic sectors (right) for Bulgaria (top) and Czech Lands/Czechoslovakia (bottom). The latter series are for Czech Lands before the First World War, the Protectorate of Bohemia and Moravia during the Second World War, and Czechoslovakia for all the remaining years. Independent in agriculture include farmers and helping family members. TKSZ and JZD are agricultural cooperatives in Bulgaria and Czechoslovakia, respectively.

Source: Own calculation based on sources detailed in Appendix.

Fig. 4 presents labor force shares by social groups and by sectors for our studied period. Bulgaria had a predominantly agricultural economy as late as the 1950s. Accordingly, independent farmers were by far the most numerous social group, comprising between 70% and 80% of the labor force before 1946. Workers consisted around 15% of the labor force, and were the only social category relatively evenly represented in all three major economic sectors. Independents in industry and services (e.g. artisans, business owners, or free professionals) and employees, practically all of which were occupied in the tertiary sector, jointly comprised the remaining 10% of the labor force.

Czechoslovakia had a significantly different socio-economic structure. Industry occupied as much as 40% of the labor force already before the First World War and its share grew over time. The share of services in the labor force also steadily increased. Naturally, the share of agriculture diminished. Workers were the largest social group throughout the studied period. They accounted for approximately half of the labor force. At the same time, the share of employees increased reaching approximately a quarter of the labor force by 1950. Most workers were engaged in industry, whereas most employees were occupied in services. Independents comprised a stable share of the labor force: 30% were independents in agriculture and 10% independents outside of agriculture.

Communism brought major changes in the socio-economic structure of the labor force in both countries. In Bulgaria, from the 1950s, the share of both industry and services started to increase at the expense of agriculture. Independent farmers, who had previously been the dominant social group, shrank to about a quarter of the labor force in 1956, and due to the progressing collectivization of agriculture soon thereafter all but disappeared. Meanwhile, a new social category, formed in large part from previously independent farmers, appeared: cooperated farmers.²² In 1960, they accounted for as much as 40% of the labor force, but their share steadily decreased thereafter. The shares of workers and employees, by contrast, steadily increased during communism.

In Czechoslovakia, workers remained by far the most numerous social category, followed by a large share of state officials and employees. During socialism the share of employees increased to around 40%. As in Bulgaria, independent farmers were by and large collectivized, which produced a large group of cooperated farmers.²³ Collectivization of agriculture and nationalization of

²² Members of labor-cooperative agricultural holdings, in Bulgarian abbreviated as TKZS.

²³ Unified agricultural cooperatives, in Czech abbreviated as JZD.

industry meant that in 1961, a meager 1% of the labor force earned a living independent of the state. Among them, remaining independent farmers were the most numerous. Thus, communism made the previously vastly different socio-economic structures in Bulgaria and Czechoslovakia more alike.

4.2. Incomes

We consulted a plethora of official statistical sources to gather information on incomes of different social groups. We managed to ascertain how incomes of social groups varied across economic sectors and over time. For workers, we were able to estimate pay-gaps by gender. Whenever possible we collected data on average annual earnings, instead of wage rates, to account for other types of income, such as payments in kind or overtime work. Where necessary, we annualized incomes relying on information on working time from national income studies. Before communism, we estimate inequality of pre-tax income, defined as income before personal income taxes and government transfers. The income concept under communism refers to post-tax income, that is, after direct taxes withdrawn at source (notably payroll taxes). However, given the generally linear schedule of direct taxes under communism, the difference between pre-tax and post-tax income inequality was minimal.

For the pre-socialist period, we compiled income data mainly from statistical yearbooks and national income studies. Bulgarian statistical yearbooks detail wages of industrial workers across many branches around 1911, in 1921, and from 1924 to 1945. Chakalov (1946) and Ivanov (2012) provide wages of workers in agriculture and services for the same years. We estimate incomes of other social groups in industry and services by applying premia relative to the worker wage. For these premia, we rely on Chakalov (1946) and Ivanov (2012) who estimated incomes of different social groups across economic sectors.²⁴

We innovatively estimate incomes of independent farmers. Simply assigning an average income to a social group that accounts for the majority of the labor force would grossly underestimate inequality. We therefore combine information from national income studies and land surveys to distribute independent farmers and their incomes into multiple income groups classified by land size. In Bulgaria, for example, starting from national income from agriculture, we deduct estimated labor income and distribute the remaining land income by income classes. To do so, we employ land distributions for multiple benchmark years. To each income class we then impute a labor income so that farmers in the lowest category of land ownership earn little beyond their imputed labor income, while farmers in higher land categories earn more, depending on the size of their plot.

Czechoslovak statistical yearbooks document wages of workers within branches of agriculture and industry, in 1913 and during the interwar period. Statistical yearbooks of the Protectorate of Bohemia and Moravia cover the period from 1939 to 1945. We estimate incomes of employees in industry and agriculture by multiplying branch-level worker wages with employee premia that we calculate based on Stádník (1946).²⁵ The employee premia is the ratio between the average remuneration of white- and blue-collars. We take salaries of public employees from statistical yearbooks. We rely on Stádník (1946) for total income of independent farmers and distribute their income following the same procedure as for Bulgaria. In the same spirit, we distributed independents in industry and services into multiple income groups based on tax statistics that reported their income from business and self-employment. Thus, we estimate the number of independents and their incomes in agriculture based on land statistics while we do the same for industry and services using tax statistics.

For the socialist period, we draw most information from statistical yearbooks, which also reported results of land and household budget surveys. Bulgarian statistical yearbooks provide detailed information on average wages and wage distributions in the state sector (wage distributions were based on the enterprise censuses). These include state enterprises and cooperatives, but exclude cooperated farmers (TKZS). We collect the reported branch-level wages of workers, assistants, and employees occupied in industry. These data provide insight into income differences of various social groups in the secondary sector. We do not observe wages by social groups in construction and in services. For these sectors, and for cooperated farmers, we differentiate average wages by income classes using data on wage distributions. We do this to account for within variation in incomes. Starting from the average wage, we estimate the ratio to the mean for five income classes: P0-30, P30-50, P50-70, P70-90, P90-100.²⁶ Average wages are available annually from 1955, while we observe wage distributions once every couple of years starting from 1960. We estimate incomes of any remaining independent farmers following the same procedure as for the pre-socialist period, taking the land distribution and the total value of agricultural production from statistical yearbooks.

In socialist Czechoslovakia, statistical yearbooks document wages of workers in agriculture and industry. We estimate the incomes of other social groups employed in the material sector, by applying premia relative to worker wages. The exception are independent and cooperated farmers. We distribute independent farmers as we did for the pre-socialist period, relying on annual land distributions and total value of private agricultural production from statistical yearbooks. For cooperated farmers and service branches within the non-material sector, as for Bulgaria, we combine branch-level average incomes and wage distributions to differentiate wages across the same five income classes.

Fig. 5 presents incomes relative to the mean by social groups and by economic sectors for our studied period. The most striking finding revealed by the figure is the compression of relative incomes in both Bulgaria and Czechoslovakia during socialism. Before 1945 there was a clear differentiation of relative incomes by social groups. On average, employees earned more than workers and

²⁴ See Appendix A.2 for details on income premia in Bulgaria.

²⁵ See Appendix A.4 for details on income premia in Czechoslovakia.

²⁶ The respective distributions in HBS reports and Enterprise censuses were reported in a tabulated form, classifying individuals into discrete income bands. We apply these tabulations in the G-pinter tool to summarize the entire distribution by g-percentiles (using generalized Pareto interpolation; <https://wid.world/gpinter/>).

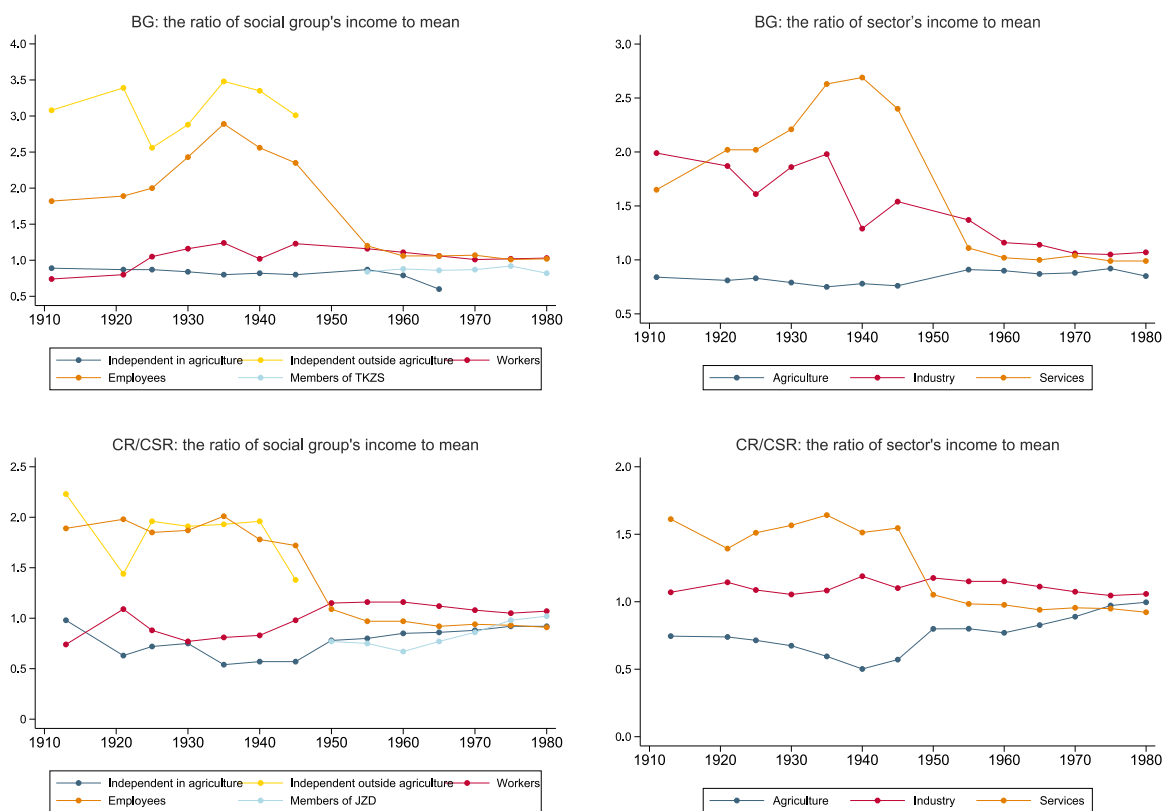


Fig. 5. Income to mean, by social groups and economic sectors.

Notes: Figures illustrate income relative to mean by social groups (left) and economic sectors (right) for Bulgaria (top) and Czech Lands/Czechoslovakia (bottom). The latter series are for Czech Lands before the First World War, the Protectorate of Bohemia and Moravia during the Second World War, and Czechoslovakia for all the remaining years. Independent in agriculture include farmers and helping family members. TKSZ and JZD are agricultural cooperatives in Bulgaria and Czechoslovakia, respectively.

Source: Own calculation based on sources detailed in [Appendix](#).

independent farmers, and self-employed in industry and services by and large earned the most. Correspondingly, incomes were the highest in services, followed by industry, and lowest in agriculture. After 1945, relative income differences between social groups diminished. The employee premia disappeared. Workers even became slightly better paid than employees. This shift was also reflected in sectoral incomes. Incomes in the material, industrial sector surpassed those in the non-material, service sector. Members of agricultural cooperatives earned similar to or even more than any remaining independent farmers.

There were other significant changes in relative income trends prior to 1945. For example, during the Great Depression, the relative income gap between farmers and other social groups widened. Farmer's incomes fell while incomes of other social groups rose. The relative gap between farmers and the rest was larger and grew faster in Bulgaria than in Czechoslovakia. This suggests increasing inequality in Bulgaria during the depression period. By the same token, falling incomes of employees and independents outside of agriculture from the mid-1930s would suggest that the Great Depression in Bulgaria was followed by falling inequality.

5. Results

5.1. Long-term income inequality from social tables

Fig. 6 presents the results on the evolution of inequality expressed by the Gini coefficient in Bulgaria and Czechoslovakia over the course of the twentieth century. The figure shows that inequality was markedly higher in Czechoslovakia than in Bulgaria in the first half of the century, with on average around 10 percentage points difference in the Gini coefficient. Inequality moved in opposite directions in the Great Depression, increasing in Bulgaria and falling in Czechoslovakia, owing to structural differences of the two economies. Inequality dramatically declined in both countries after the Second World War and the communist accession to power. The Gini coefficient fell in Bulgaria by almost 15 percentage points relative to its average pre-war levels, while it more than halved in Czechoslovakia, decreasing by more than 20 Gini points. The post-war compression was practically immediate, leading to a synchronization of inequality in the two countries by the early 1950s, with the Gini coefficient broadly converging at levels around 20 in the post-war decades. Thus, our main long-term results reveal that the two studied countries previously displaying markedly

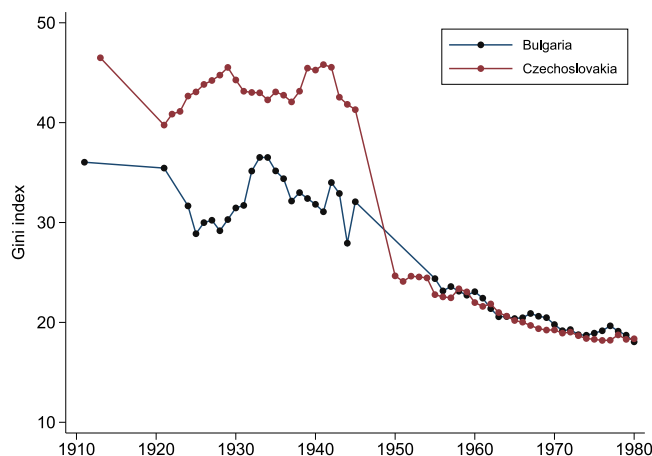


Fig. 6. Gini coefficient in Bulgaria and Czechoslovakia, 1910–1980.

Notes: The Gini series is for Czech Lands before the First World War, the Protectorate of Bohemia and Moravia during the Second World War, and Czechoslovakia for all the remaining years.

Source: Own construction based on sources detailed in Appendix.

different inequality levels suddenly became characterized by the most egalitarian distribution of income in the post-Second World War era.²⁷

Considering the inequality evolution of each country, inequality in Bulgaria declined more noticeably in the early 1920s, when the Gini fell from levels around 35–36 in 1911/1921 to 29 by 1925. There was, however, a sharp increase in inequality during the Great Depression, which signified a peak of inequality in Bulgaria (with Gini reaching 37 in 1933/4). Inequality then declined in the late 1930s, but again increased during the Second World War (seeing a sharp drop in 1944). Income inequality was dramatically reduced by the time our series reemerges in 1955, with the recorded Gini index at 23. The remaining period up to 1980 saw a further steady decline in inequality.

In Czechoslovakia, we observe a more marked decline in inequality after the First World War. Gini fell from levels around 46 in 1913 (in the Czech Lands) to 40 in 1921 (in Czechoslovakia). Inequality gradually increased by the end of the 1920s, when Gini reached 45 in 1929. The Great Depression in Czechoslovakia was featured by a decline in inequality (the Gini decreased to 41 by 1934), and it was followed by a moderate increase in the late 1930s. The Gini coefficient assumed higher levels in the newly established Protectorate of Bohemia and Moravia, but decreased during its existence from 46 in 1939 to 41 in 1945. By 1950, the Gini dramatically decreased by 15 Gini points, and the decline continued during the whole next decade when it stabilized at low levels slightly below 20.

5.1.1. Social transformations

We investigate in more detail documented inequality trends in the two countries by applying the social class analysis. As already mentioned, our dataset is unique given the level of detail about social classes in terms of their relative size and income, but also about their inner stratification. Hence, it is possible to consistently track the secular developments in society's social stratification and to ascertain how social changes impacted inequality patterns. We proceed in this direction by presenting stacked kernel densities by social classes at different time snapshots throughout the twentieth century (Jenkins, 1995; Lakner and Milanovic, 2016).

Bulgaria. We first look at the development of kernel densities by social groups in Bulgaria at four snapshots at similar time intervals – specifically in 1911, 1930, 1955, and 1970 (Fig. 7).²⁸ It can be seen that the income distribution in Bulgaria in the first half of the twentieth century (in 1911 and 1930) was characterized by the overwhelming concentration of the population around its (unique) mode in the middle of the distribution. In addition, the kernel density is slightly positively skewed with a relatively smaller concentration of individuals in the upper-middle income levels. The figure further reveals that sectoral development played a key role in shaping inequality in the first half of the twentieth century. Namely, density plots for 1911 and 1930 show that the independent farmers made up the bulk of the population (almost three-fourths of the population; the blue area), which coupled with the markedly

²⁷ It is sometimes argued that income inequality during communism was actually (significantly) higher because statistical sources (such as household surveys) did not usually include the communist elites, the so-called *nomenklatura*. It is, however, unlikely that the inclusion of *nomenklatura* would overturn the general picture of marked equality of income during the socialist period. First, the communist *nomenklatura* was by definition a rather small (exclusive) group, which implies that its impact on overall inequality was limited (i.e. when measuring inequality among all individuals in the country), and this holds, in particular, for measures such as the Gini coefficient which are less sensitive to changes at the distributional tails. Second, direct evidence from personal income tax data on high-income individuals during communism suggest relatively low top concentration (for example, lower than in contemporary western countries; see Novokmet (2017) for socialist Yugoslavia, Bukowski and Novokmet (2021) for socialist Poland). Furthermore, even the “corrections” for the non-monetary privileges of the communist elites (Morrison, 1984; Atkinson and Micklewright, 1992) did not alter the general finding of very low income inequality during socialism.

²⁸ We look at 1955, which is the first year after the Second World War in our dataset for Bulgaria.

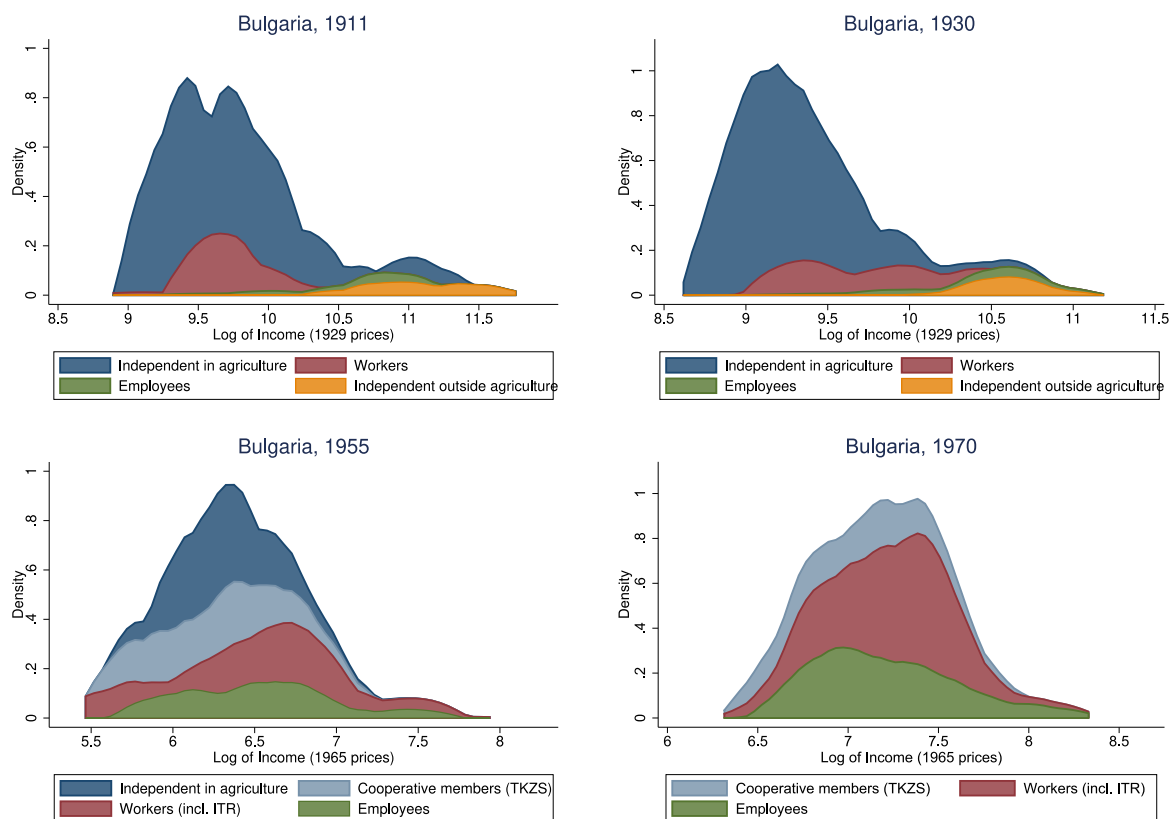


Fig. 7. Size and relative income of social groups in Bulgaria in 1911, 1930, 1955, and 1970.

Notes: Independent in agriculture include farmers and helping family members. TKSZ are agricultural cooperatives in Bulgaria. ITR stands for engineering and technical workers. Kernel densities are estimated before imputing within-inequality based on the log-normal distribution (see Table 1).

Source: Own calculation based on sources detailed in Appendix.

egalitarian distribution of land ownership entailed rather moderate dispersion of income in the agricultural sector. The two top panels of the figure show that the size of this group remained relatively constant in the first half of the century (thus suggesting the absence of structural change), while the concentration within this subgroup saw moderate changes. The most material change in the density curve for the independent in agriculture between 1911 and 1930 consisted in their virtual disappearance from the right tail and a marked shift towards the middle of the distribution. This was likely the outcome of the land reform under the agrarian government of Alexander Stamboliyski in the early 1920s, which targeted the few large estates in the country (Bell, 1977).

Turning to other groups, the mass of workers, as the second largest social group in pre-communist Bulgaria (accounting, however, for less than 15% of the labor force; compare with Fig. 4) resided in the (lower-) middle parts of the distribution, while the densities of employees and independents outside of agriculture were largely concentrated in the right-hand part of the distribution. Moreover, the figure reveals that the relative distance between these groups and independents in agriculture was the chief determinant of inequality. Put differently, given the distinct character of these groups, we can say that the urban-rural gap was the main source of inequality. Thus, when the densities of the groups shifted apart more substantially by 1930 – and the right tail elongated – income inequality expressed by the Gini coefficient increased (Fig. 6). This conforms with the peak of inequality during the Great Depression when the urban-rural gap surged as a result of the so-called “price scissors”.^{29 30}

The two bottom panels of Fig. 7 suggest that a transition to the communist system in Bulgaria after the Second World War was truly a radical social revolution. As already mentioned, the communist government forced the “big-push” industrialization by adopting a strategy that internalized basic tenets of the dual economy model, relying on the massive labor reallocation from agriculture to industry. This development entailed profound social transformation, which primarily and fundamentally affected

²⁹ The episode of “price scissors” is visible from the evolution of terms of trade and the ratio of industrial to agricultural prices. A sharp drop in international agricultural prices during the Great Depression hit especially hard the net exporters of agricultural products (importers of industrial products), in the first place the countries in South Eastern Europe. The overreliance on agricultural exports (imports of industrial goods) made these countries excessively vulnerable to shocks in international prices. See, for example, Lampe and Jackson (1982) on Bulgaria and Olšovský (1961), Prucha (2004) on Czechoslovakia.

³⁰ Even an increase in dispersion among workers between 1911 and 1930 was primarily due to an enlarged income gap between agricultural workers and other (non-agricultural) workers by 1930.

independent farmers, who were either compelled to join the collective farms (TKZS in Bulgaria; modeled after the Soviet kolkhozes) or migrate to urban areas with rapidly expanding industry (thus also alleviating hidden unemployment in agriculture, which had assumed endemic levels in the interwar era). As the bottom left panel of Fig. 7 suggests, the share of TKZS members in the labor force reached around 30% by 1955 already (see also Fig. 4).³¹ Forced collectivization of agriculture allowed communists to squeeze resources from agriculture and use them for investment in industry (Nove, 1961; Lampe, 1986).³²

At the same time, the groups of workers and employees expanded with the rapidly growing secondary and tertiary sectors. Workers became the most numerous social group in Bulgaria by the mid-1960s. Interestingly, the kernel density plot for 1955 in Fig. 7 suggests a still relatively dispersed within-distribution of workers (relative to 1930).³³ It may be conjectured that this was due to a still more considerable inequality between industrial branches in the early socialist period (for example due to more marked wage discrimination between the so-called “producer” and “consumer” goods industries),³⁴ as well as indicative of the more pronounced skill-premium commanded by engineers and technical personnel (ITR in Bulgaria) in the early phases of the industrialization with the large supply of labor migrating from the rural sector. But as industrialization advanced and the share of workers continued to expand, their within-distribution became more compressed (especially once labor shortages kicked in). The density plot for 1970 reveals that an enlarging mass of workers heaped in the middle parts of the distribution, while both tails shortened (the bottom right panel in Fig. 7).

Finally, the long-run trends for the two top classes in the pre-communist income hierarchy – independents outside agriculture and employees – further corroborate that the introduction of communism thoroughly transformed patterns of social stratification in Bulgaria. The independents outside agriculture were liquidated as a distinct social class, while the nature of the white-collar occupations, together with their social standing and relative income, was fundamentally transformed. Given that in terms of relative incomes these two groups exhibited the largest distance to the rest of the population (or “premium”) in pre-communist Bulgaria (see also Fig. 5), their disproportionate material depletion after the Second World War critically impacted overall income inequality (by reducing the between component of inequality; see below). And since both groups were in the aggregate more important in Czechoslovakia before communism, we appraise their fortune in more detail when we next analyze the density plots for Czechoslovakia.

Czechoslovakia. Fig. 8 reports the kernel densities for Czechoslovakia by taking the same twenty-year snapshots – namely in 1913, 1930, 1955 and 1970. The upper two panels of the figure show that the social structure in Czechoslovakia in the first half of the twentieth century differed markedly from that in contemporary Bulgaria. Correspondingly, the density curve for pre-war Czechoslovakia was more dispersed, with a fatter right tail.

Pre-communist Czechoslovakia presented a clearly demarcated class society broadly comparable to the industrial societies of contemporary Western countries. In a country that was already considerably industrialized by the early twentieth century (see Fig. 2), workers accounted for the largest social group. They were largely concentrated in the lower-middle and middle parts of the income distribution. The critical factors for the dispersion of workers was the sector of occupation (concretely, agriculture versus the rest) and gender (see below).

Next, the social standing of employees was notably advantageous in Czechoslovakia before the Second World War. This was a broader Central European phenomenon, where the privileged status of salaried employees was institutionalized (such as in the separate insurance scheme and various legal provisions from job security to benefit schemes; originating from Germany and Austria–Hungary (Kocka, 1981)). The considerable income premium earned by employees relative to workers (see Fig. 5) reflected to an important extent the “status” premium (Kalecki, 1993; Scitovsky, 1966).³⁵ A material basis of this social differentiation was further strengthened during the First Czechoslovak Republic, as evidenced, for example, by an increase in the employee premium (with the introduction of new salary scales in 1926 that particularly favored civil servants (Teichova, 1988)).

An important dimension that underlay the earning distribution of both workers and employees was gender. Pre-communist Czechoslovakia was characterized by high gender inequality in earnings.³⁶ The bottom left panel of Fig. 9 shows clear-cut segregation by gender of workers and employees in Czechoslovakia in 1930. Densities of both female workers and employees were concentrated at the lower income levels, noticeably left of the main mass of their fellow male workers and employees. High gender inequality in earnings was due to several factors. First, female workers were disproportionally employed in a handful of industries – notably in agriculture, textiles, clothing or food processing – which paid below average wages (agriculture in particular, as mentioned above; according to the 1930 census, almost 40% of female workers were employed in agriculture). However, industrial segregation by

³¹ Collectivization of agriculture was the most expedient in Bulgaria as compared to other socialist countries in Eastern Europe.

³² This reasoning followed the logic of “primitive socialist accumulation”, as Preobrazhensky famously postulated (on the Soviet industrialization debate, see Dobb (2012); Erlich (1960); Allen (2003)).

³³ Measures of within-inequality among workers, such as the Gini index or the mean log deviation, display a relatively modest decline of inequality. For example, the Gini index for workers was 31.2 in 1930, while 27.6 in 1955.

³⁴ Also reflected in a higher wage gap between workers in heavy industry or construction, on the one hand, and workers in state-owned agriculture, on the other.

³⁵ Kocka (1981) suggests a sharp (and “socially relevant”) delineation of salaried employees and wage earners (*Angestellte* and *Arbeiter*) in Germany or Austria in the late nineteenth and early twentieth centuries. Moreover, he suggests an antagonistic stance of employees towards workers.

³⁶ We analyze gender inequality among workers and employees only, because the statistical sources only allow us to distinguish industry (branch) specific earnings by gender. Although dictated by the data, this focus is conceptually sensible – and hence assumed in the literature – as male and female workers/employees have generally been engaged as independent individuals in the labor market. On the other hand, other social groups, such as independent or cooperated farmers, have generally generated income together as a household; accordingly, we have equally split income of a household between constituent adults (see Section 3 for details about the population units of analysis).

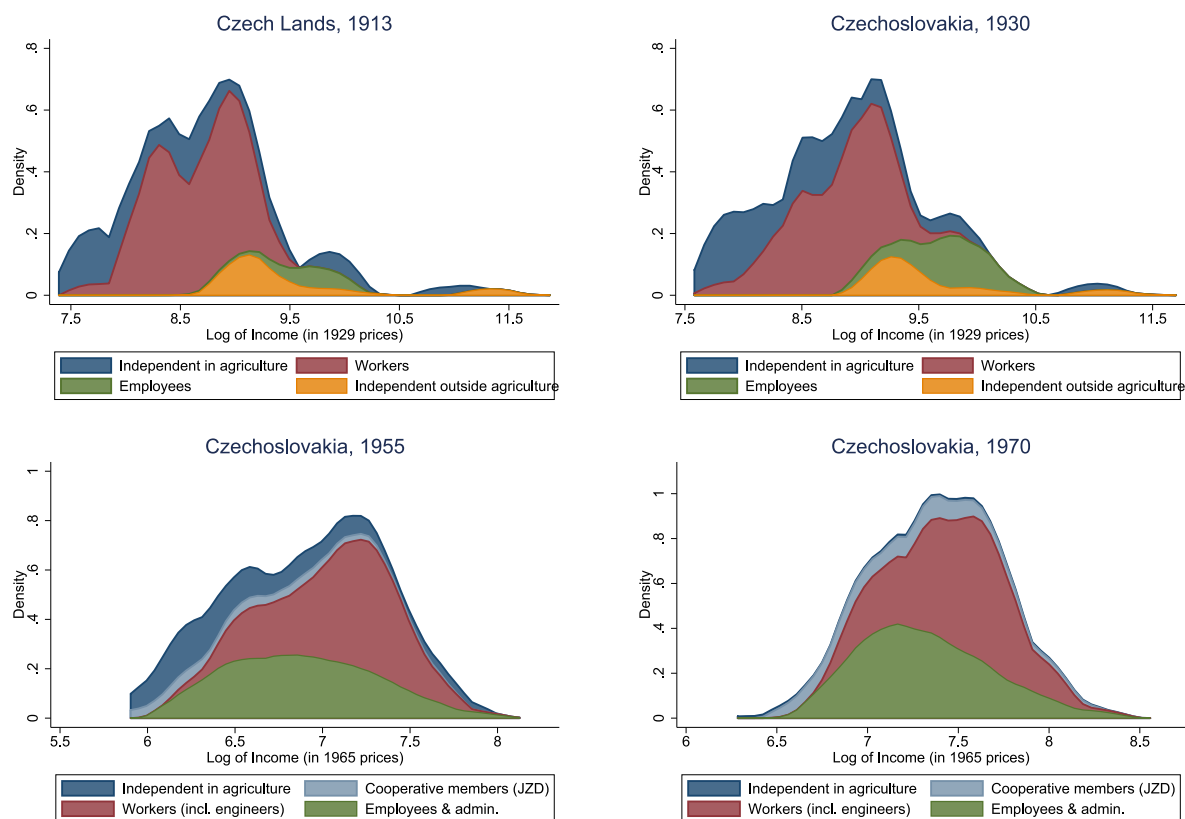


Fig. 8. Size and relative income of social groups in Czech Lands/Czechoslovakia in 1913, 1930, 1950, and 1970.

Notes: Independent in agriculture include farmers and helping family members. JZD are agricultural cooperatives in Czechoslovakia. Kernel densities are estimated before imputing within-inequality based on the log-normal distribution (see Table 1).

Source: Own calculation based on sources detailed in Appendix.

gender seems to be only a (smaller) part of the explanation, as we observe the pronounced within-industry gender gap in earnings. Explaining this is beyond the scope of this paper, but it is plausible to attribute it to differences in skill requirements of jobs assumed by males and females (i.e., females generally had low-skilled jobs within a specific industry/sector due to shorter job experience, limited access to apprenticeship, fewer working hours, etc. Mesch (1984)).³⁷

Lastly, independents outside of agriculture made a heterogeneous group. As the kernel density plots for 1913 and 1930 reveal, a greater mass of them was concentrated in the mid-income ranges, and could be treated as a part of the middle classes – probably largely comprised of self-employed such as artisans, craftsmen, shopkeepers, or small entrepreneurs (Teichova, 1988). This group, however, also comprised true capitalists, such as large industrialists or financiers, who populated the long right tail of the distribution. One could count among this income super-elite the richest landlords who were predominantly engaged in commercial agriculture (included in the category of independent in agriculture). We look below in more detail at this upper crust of the pre-communist Czechoslovak society (Section 5.2).

Communist Czechoslovakia was in the true sense of the word the laboratory of social change (Krejčí, 1972; Krejčí and Machonin, 1998). Profound social transformation accompanied the introduction of communism, which involved a gigantic equalization between social classes. The bottom left panel of Fig. 8 shows that the elongated right tail, that defined the pre-1945 distribution of income, was thoroughly chopped off by the 1950s. Both large capitalists immediately after the Second World War,³⁸ and of all other independents outside agriculture soon thereafter, were expropriated by communist nationalizations and expropriations. Big land reforms deprived large (and medium-large) landholders of their landholdings, while the remaining independent in agriculture were gradually pressured to join agricultural cooperatives (JZD) as the process of land collectivization progressed.

³⁷ Mesch (1984) provides a comprehensive overview of gender inequality in earnings among workers in Imperial Austria before the First World War. He explains the concentration of women in low-skilled jobs by the fact that female industrial workers faced challenges such as limited education, societal devaluation of formal training, and employer reluctance to invest in female education due to short job tenures. Furthermore, institutional barriers confined women to less physically demanding but monotonous tasks. Short job durations, influenced by age and marital status, hindered career growth, perpetuating wage disparities and underrepresentation in unions.

³⁸ Note that considerable nationalization of industry had occurred even before communists' accession to power (e.g. with Beneš decrees).

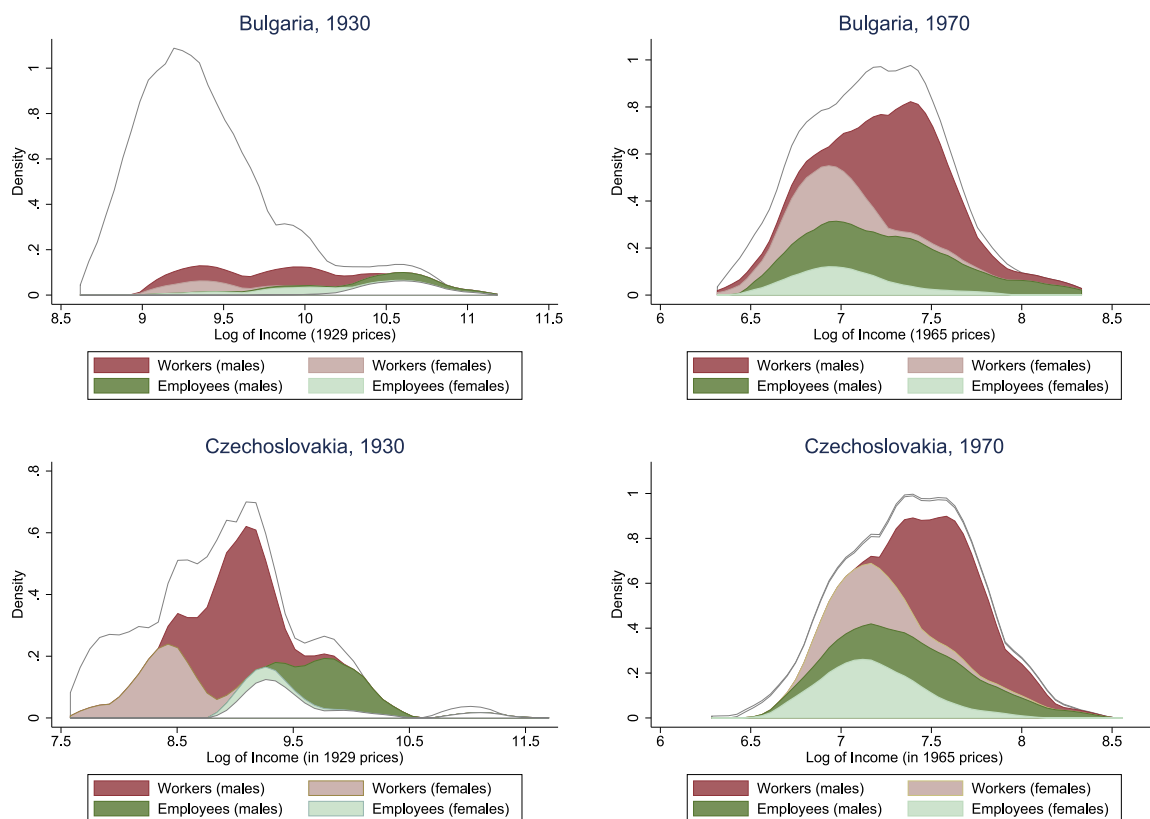


Fig. 9. Size and relative income of workers and employees by gender.

Notes: Kernel densities are estimated before imputing within-inequality based on the log-normal distribution (see Table 1). The blank densities in the chart refer to independent in agriculture and independent outside agriculture (as in Figs. 7 and 8).

Source: Own calculation based on sources detailed in Appendix.

Next, the material, and consequently social, standing of employees was considerably devalued relative to the pre-Second World War period (Maňák, 1967). The separate insurance scheme was abolished and all other legal privileges enjoyed before the war removed (Krejčí, 1972). It should be mentioned that the process of relative deterioration of the employees' position was already initiated during the German occupation (both to target national intelligentsia and stimulate workers for the war effort; Krejčí and Machonin (1998)). During the socialist period, the relatively advantageous standing of workers vis-à-vis employees was more pronounced in Czechoslovakia (as conveyed by the lower two panels of Fig. 8; see also Fig. 5). Formal remuneration structure rewarded more physically intensive jobs, such as in heavy industry or construction, relative to intellectual jobs, largely concentrated in the tertiary sectors with employees (Večerník, 1991; Flemming and Micklewright, 2000).

Communism also brought about a reduction in gender inequality in earnings (the bottom right panel of Fig. 9). Equality between sexes was one of the basic tenets of communist ideology (it was formally incorporated in the 1960 Czechoslovak Constitution and in the labor legislation).³⁹ Improvements in the relative standing of women were additionally driven by significant advancements in their educational attainments and the secular shift of female employment from low-wage sectors (especially working in agriculture and as domestic servants). Notwithstanding these advancements, Fig. 9 shows that the actual gender earnings gap in communist Czechoslovakia was far from the proclaimed ideal. Moreover, it was often argued that gender was the most important dimension for overall income inequality in the otherwise distinctly egalitarian society (Večerník, 1991; Atkinson and Micklewright, 1992). We have already indicated some of the factors contributing to tangible gender differences in earnings. The gender gap between workers was to an important degree a result of industrial segregation of employment, where women were heavily underrepresented in the highest-paying sectors, such as in heavy industry or construction (overrepresented in the low-paying, such as trade and catering, or textiles (Michal, 1973)). As the overall placement of the density plot for male workers in the bottom right panel of Fig. 9 shows, this new "labor aristocracy" of industrial working men was among the highest-paid groups in Czechoslovakia. Furthermore, women were disproportionately employed in the tertiary sectors generally paying below average earnings. The density plots for 1970 also

³⁹ For example, according to the Law (No. 65/1965), women were given equal rights with men in the workplace, as well as guaranteed equality in the area of wages (Havelková, 2009, p. 192).

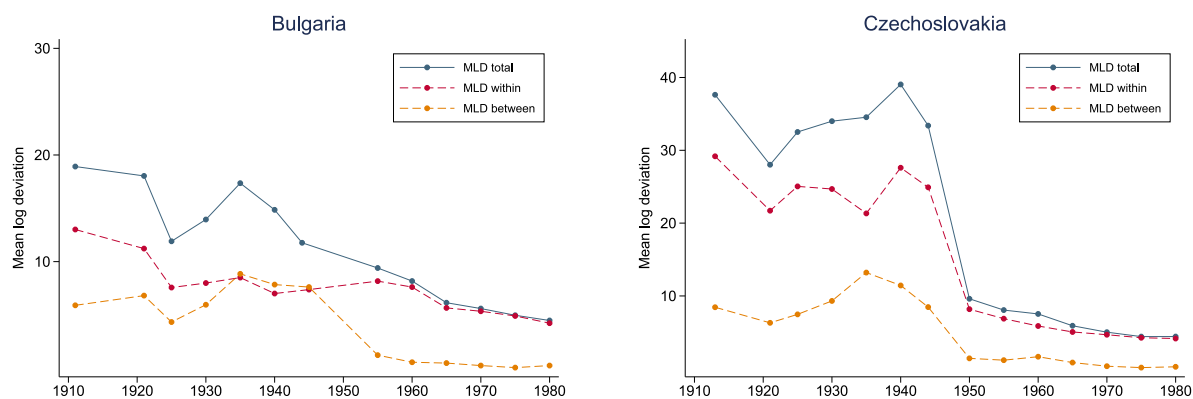


Fig. 10. Mean log deviation (MLD) of income: overall, between and within social groups.

Source: Own calculation based on sources detailed in [Appendix](#).

reveals the existence of the non-negligible gender gap among employees, which can be reconciled with historical accounts pointing to a marked male dominance in high-paying administrative jobs.

Finally, it may be pointed to the primacy of ideological considerations in driving the social transformation in Czechoslovakia. Namely, practical concerns, notably industrialization demands were less pressing in industrialized Czechoslovakia.⁴⁰

5.1.2. Between and within decomposition

Density plots have provided broad intuition of how developments in the relative size and incomes of social groups impacted trends in income inequality during the twentieth century. Equipped with these insights, we proceed by considering a formal decomposition of inequality into inequality between and within social groups. For this purpose, we decompose Mean Log Deviation (MLD), which in contrast to the Gini index, is exactly additively decomposable measure of inequality by population subgroups. That is, the total inequality measured by MLD may be expressed as the (population-weighted) sum of inequality between the subgroup mean income and inequality within subgroups.

Fig. 10 presents the results of the MLD decomposition in Bulgaria and Czechoslovakia over the study period. First, it can be seen that the development of overall inequality as measured by MLD indicates the same trends in inequality as suggested by the Gini index above (Fig. 6). Notably, both measures suggest a secular decrease in inequality occurring with the communist accession to power after WW2.

Looking at the contribution of the within and between components to the overall income inequality, it can be seen that the within-inequality component was a relatively more important source of inequality in both countries at the beginning of the twentieth century. This was particularly the case in the Czech Lands before the First World War, where the within component accounted for almost 80% of inequality in 1913. Accordingly, a reduction in the importance of the within component after the war was the main factor behind a noticeable decrease in overall inequality in the Czech Lands between 1913 and 1921. As mentioned above, inequality among independents in agriculture decreased after the Great War in both Bulgaria and especially Czechoslovakia due in part to the major land reforms.⁴¹ More importantly, a decline in income inequality among independents in agriculture also reflected the “return” component, that is, a decline in agricultural prices in the early 1920s relative to the pre-First World War period disproportionately hurt incomes of larger landholders.⁴²

Plausibly, the most important reason for a decline in the between component in Czechoslovakia after the First World War was a marked increase in the relative incomes of workers (Fig. 5), who were the most numerous group in the country. The reasons for this noticeable improvement in workers’ standing should be sought in the immediate post-war development that armed workers had higher bargaining power, and in turn pushed for wage increases.⁴³ At the same time, the relative income of employers (independents outside agriculture) was considerably reduced and reached in the early 1920s its interwar low (Fig. 5).

⁴⁰ This was embedded in the very nature of the new state. Teichova (1988, p.157) refers to the 1960 constitution (according to which Czechoslovakia was officially declared a socialist state): “It was laid down that this state was based ‘on the firm alliance of workers, peasants and the intelligentsia [...] led by the working class’”.

⁴¹ In Czechoslovakia, the pre-war land inequality was very high, and the extent of the redistribution more extensive as a result. Bulgaria was, by contrast, already characterized by the relatively egalitarian land distribution before the war, and the extent of the land redistribution – and a resulting decrease in the within component of inequality – was less pronounced (as mentioned above, and can be seen in Fig. 10 (left panel), the strongest decline in within inequality of independents in agriculture occurred in the early 1920s).

⁴² More generally, we see that the periods of agricultural booms (busts) were accompanied by the rise (fall) in inequality among independent in agriculture. In Czechoslovakia, the agricultural boom of the late 1920s disproportionately benefited middle and large landholders specialized in commercial crops (inequality among independents increased), while the Great Depression had an opposite effect (inequality among independents decreased).

⁴³ Sometimes claimed as excessive, for example as discussed for Germany under the so-called Borchardt controversy.

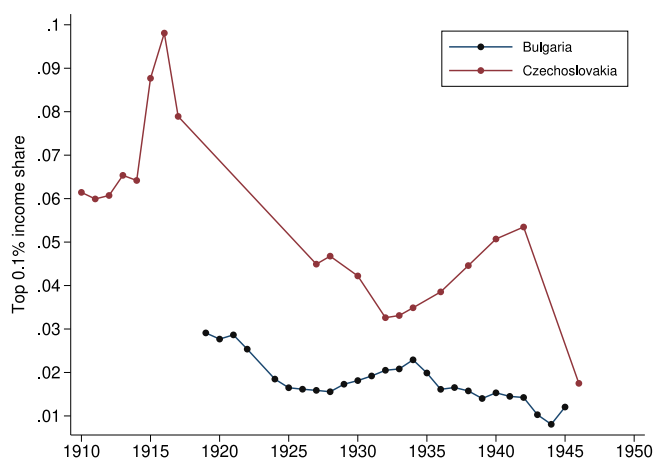


Fig. 11. Top 0.1% income share in Bulgaria and Czech Lands/Czechoslovakia, 1910–1946. Source: The series for Czechoslovakia are from Novokmet (2023); Bulgaria: own construction.

Both countries saw an increase in the between component during the Great Depression, but its impact on the overall inequality differed. In Bulgaria, the between component drove the overall increase in inequality, which is consistent with the above-expounded upsurge in the urban-rural income gap in this period (also suggested above by the development in mean incomes in Fig. 5). In Czechoslovakia, in contrast, a decline in the within component (among independents inside and outside agriculture) outweighed a rise in the between component, with the net effect being the decline in overall inequality during the Great Depression.

The secular compression after the Second World War was driven by a decreasing contribution of both components. However, the importance of the within component decreased only modestly in Bulgaria, primarily because the dominant group of independents in agriculture had already displayed relatively low within inequality before the war. Specifically, the social transformation in the post-Second World War decades in Bulgaria entailed rising population shares of workers and employees which displayed similar within inequality as independents in agriculture before the war. On the other hand, the importance of the within component fell more significantly in Czechoslovakia after the Second World War, induced by a decrease in inequality among all social groups.⁴⁴

During the socialist era, inequality became almost exclusively determined by the within component, itself reduced to quite moderate levels. As Fig. 10 shows for both countries, the importance of the between component for inequality practically vanished, which, at least in terms of income hierarchies between social classes, came closest to the communist ideal of a classless society.

5.2. Top income shares

We mentioned in Section 2 that Kuznets (1955) also proposed the alternative mechanism of concentration of saving among high-income groups as inducing an increase in inequality along the development path. The evidence on top income shares could be informative to ascertain this mechanism.

The historical accounts suggest that the industrialization of today's developed countries led to the growing concentration of capital income and rising top income shares. Allen (2009), for example, argues that the British Industrial Revolution was initially accompanied by rising profits amid stagnant wages (the so-called “Engels pause”).⁴⁵ It was also mentioned above that the dual-economy model postulates a rising capital share as the modern capitalist sector expands amid structural change (Ravallion, 2016). This development implied – with capital income being strongly concentrated – that the gains from the growth take-off largely went to high incomes. And given that high incomes have a higher marginal propensity to save, there was a tendency of further wealth and capital income concentration.⁴⁶

Kuznets' conjecture of cumulative effects of concentration of saving on top income shares provides a useful framework to compare top income shares in Bulgaria and Czechoslovakia from a development perspective. Fig. 11 shows the development of the top-0.1% share in the two countries in the first half of the twentieth century. It can be seen that the very top income shares were markedly higher in Czechoslovakia and experienced notable volatility relative to their Bulgarian counterpart.

⁴⁴ This is consistent with trends documented in Novokmet (2023), Fig. 9, which show a strong decline in inequality for both wage earners and salaried in Czechoslovakia after the Second World War.

⁴⁵ That is there was a rise in the share of capital income (a decline in the share of labor income) in national income, which ensured necessary capital accumulation to sustain industrialization. Dumke (1991) suggests the same pattern for the industrialization of Germany.

⁴⁶ This logic led to the formulation of the “classical” channel positing positive effects of inequality on economic growth via capital accumulation (Keynes (1920), Kaldor (1957), Galor and Moav (2004); Itskhoki and Moll (2019) have recently argued that the optimal policy along the development path is to initially stimulate the rise in business profits (by suppression of wages) to allow higher capital accumulation).

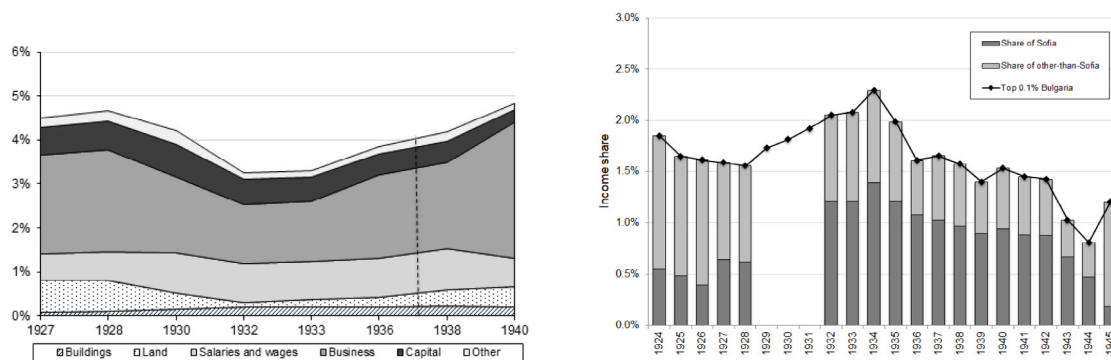


Fig. 12. Top-0.1% share decompositions in Czechoslovakia (left) and Bulgaria (right).

Notes: Left panel: Decomposition of the top-0.1% in Czechoslovakia by income sources. Right panel: Decomposition of the top-0.1% in Bulgaria between the share of Sofia and other-than Sofia.

Source: Left panel from Novokmet (2023), Figure 4b. Right panel: own construction.

Both phenomena may be explained by the strong concentration of capital income at the top of the income distribution in industrialized Czechoslovakia. The left panel of Fig. 12 clearly shows that the top-0.1% in interwar Czechoslovakia was predominantly composed of capital income – in particular, of business profits – whose dominant industrial character has been documented by Novokmet (2023). Similarly, it can be seen that shifts in business profits were the main reason behind more pronounced volatility in Czechoslovak top income shares. This is consistent with findings in the literature that explain high levels of top income shares in advanced countries in the first half of the twentieth century by the strong concentration of capital income (Atkinson and Piketty, 2007, 2010).

There is unfortunately no data on income sources of top incomes in Bulgaria, but it is tempting to attribute both substantially lower top income shares levels and their stability to the pre-industrial character of interwar Bulgaria. To put it simply, the great industrial or financial fortunes associated with the advancement of capitalism were generally lacking in Bulgaria. The available geographical distribution of the top-0.1% share on the right panel of Fig. 12 may help us to illuminate the character of Bulgaria's rich in the interwar period. In particular, a noticeable geographical rebalancing of top-0.1% during the Great Depression from traditional commercial centers in agricultural goods (such as Plovdiv, Ruse, Varna or Burgas) in favor of the country's capital, Sofia (whose share of the top-0.1% income almost doubles in the 1930s) may suggest a shift in the dominant outlook of the rich – specifically, the replacement of “merchant” with “state” capitalists.⁴⁷ Both types may be seen as more pertinent to pre-industrial stages of capitalism (e.g. Braudel (1979)), and the historical evidence consistently point to the lack of the (western-style) industrial and financial *haute bourgeoisie* in pre-communist Bulgaria.

Finally, although we believe that the industrialization axis proved to be a useful comparison framework to understand top incomes trajectories in pre-communist Czechoslovakia and Bulgaria, it is important to stress that the top inequality dynamics is critically determined in the broader institutional setting. Notably, industrialization is not necessarily accompanied by the rising wealth and capital income concentration (Roine and Waldenström, 2015), which especially hinged on historical institutions governing the land distribution.⁴⁸ Also in this respect, Bulgaria and Czechoslovakia provide a revealing contrast, with markedly different land inequality – unusually low in the former and very high in the latter – and (resultingly) different social relations in agriculture, which in part implied a distinctive industrialization path and inequality development in the Czech Lands (and more generally in Central Europe) or its absence in Bulgaria (Gerschenkron, 1962; Palaret, 1997).⁴⁹ Finally, we should underline the critical role of institutional arrangements of property relations for the inequality dynamics (Piketty, 2020; Novokmet, 2023). Concretely, a sweeping transformation from private to public capital amid communist nationalizations of private capital and the fact that capital accumulation during communism was altogether assumed by the public sector had a decisive impact on private capital income concentration in both countries.

⁴⁷ With the end of the agricultural export boom of the 1920s, Sofia assumed an excessively large role in Bulgaria's economy in the 1930s, when the traditionally large role of the state in the economy was taken one step further. Lampe and Jackson (1982, p.240) thus assert that “previous economic centers suffered [...] for being far from the seat of political power”. Ivanov and Ganey (2014) argue that the geographical rebalancing of top incomes was the result of industry moving from the countryside to the capital.

⁴⁸ High levels of wealth inequality in the pre-industrial period were typically the function of high land inequality; in these scenarios, the falling importance of agricultural land in the total wealth could have entailed a stagnation or decrease of wealth inequality during industrialization (Piketty and Zucman, 2014; Waldenström, 2023).

⁴⁹ An interesting debate considered to what extent the so-called *chiflik* estates of the late Ottoman period (see generally İnalçık (1983)) were similar to commercial demesnes in Central Europe (such as the Prussian *Gutscherrschaft*) in their assumed export orientation, bonded labor or capital accumulation. Correspondingly, it was discussed whether *chifliks* could have altered class and property relations in agriculture and stimulated industrial development (e.g., as Brenner (1982) proposed for England). See Lampe and Jackson (1982) on a negative consensual assessment.

6. International comparisons

We next turn to international comparisons. Historical social tables provide a natural starting point for comparison, specifically for Eastern European countries during the first half of the twentieth century. As already mentioned, Lindert and Nafziger's (2014) study of income inequality in Russia in 1904 is closest to ours in terms of data and methodology used to construct social tables, in particular in their pioneering combination of different data sources to estimate inequality within social classes.⁵⁰ Their study suggests moderately high levels of inequality (by standards of the time) in Russia at the beginning of the twentieth century. Income inequality expressed by the Gini coefficient was 36.2, which is remarkably similar to our estimates for prewar Bulgaria (with Gini equal to 36 in 1911). Relatively lower inequality levels can be explained by shared structural features of the two countries before the First World War: (i) both Bulgaria and Russia were predominately agricultural countries, with more than two-thirds of population employed in agriculture; (ii) the predominantly peasant population was characterized in both countries by the relatively egalitarian distribution of income.⁵¹

More generally, the same logic may be extended to explain inequality dynamics in other countries in Eastern Europe in the first half of the twentieth century, all broadly similar in terms of sectoral development and land ownership patterns (that is, generally corresponding to Kuznets' conception of the initial conditions characterizing developing countries in the dual-economy framework; see Section 2). For example, a study of interwar Yugoslavia by Vinski (1967) corroborates these conclusions.⁵² Vinski's estimates of income inequality in Yugoslavia in 1938 are quite close to levels we obtain in Bulgaria (Gini in Yugoslavia was 37 vs. 34 in Bulgaria). Importantly, he argues that income inequality in Yugoslavia was critically held down by the relatively egalitarian agricultural sector (Gini for independent farmers is 19 in Yugoslavia vs. 18 in Bulgaria). Another example is the seminal study by Wiśniewski (1934) for Poland that also explains relatively moderate inequality levels in 1929 (the Gini index equal to 34) by the relatively equal distribution of the peasant population (see also Bukowski and Novokmet (2021)).

We can also compare our series to inequality developments in industrial countries based on the annual ("dynamic") social tables recently constructed for Germany and the United Kingdom (Gómez León and De Jong, 2019). According to this study, at the beginning of the twentieth century the Gini coefficient was 43 in Germany and 45 in the U.K. These levels are very close to our estimates for the Czech Lands before the First World War.⁵³ Equally, the whole interwar period and Second World War were featured by high income inequality in both Czechoslovakia and other industrial countries.⁵⁴ We believe that these comparisons also corroborate the importance of the level of development of social structures for inequality. We saw that pre-communist Czechoslovakia was a stratified class society very much alike the industrial societies of contemporary Western countries. The shared inequality experience of industrialized Czechoslovakia with Western developed countries in the first half of the twentieth century is also confirmed by the development of top income shares (Novokmet, 2023).

It was only after the Second World War and the institution of the communist system in the country that Czechoslovak inequality diverged fundamentally from the Western European patterns. The secular decline of inequality occurred everywhere in the post-Second World War decades (e.g. Piketty, 2001; Atkinson and Piketty, 2007; Atkinson et al. 2011), but it was especially pronounced in the newly turned communist countries. This was principally due to more extreme and sweeping institutional and political changes that accompanied the introduction of communism. Accordingly, and as already mentioned, the Czechoslovak inequality trajectory might be seen as an extreme version of the inequality development in Western countries. At the same time, the transition to communism led to the synchronization of inequality in Bulgaria and Czechoslovakia, which converged to almost identical inequality levels (Fig. 6). The social structure of the two countries, and its (political-institutional) determinants, became practically indistinguishable (Figs. 7 and 8).

Finally, it is interesting to consider how income inequality evolved after the fall of communism in the two analyzed countries. For this, we rely on the official measures of inequality based on household budget surveys (HBS), instated in Bulgaria and Czechoslovakia in the late 1950s/early 1960s and conducted (more or less) regularly since then.⁵⁵ In Fig. 13 we plot them together with our

⁵⁰ Such as using information on the distribution of land to estimate inequality between independent in agriculture or using fiscal data to estimate high business and professional incomes of self-employed.

⁵¹ It should be noted that relatively lower income inequality among peasants in Tsarist Russia was primarily a consequence of relatively egalitarian access to communal land, rather than a result of the relatively egalitarian distribution of private ownership of land Lindert and Nafziger (2014).

⁵² Interwar Yugoslavia is a particularly interesting comparative case study given its quite similar sectoral structure to that of Bulgaria at the time (see Fig. 2). Vinski's study also used a social tables approach to estimate income inequality. It provides an overview of social groups in interwar Yugoslavia, importantly also allowing for within inequality among independent farmers based on the size of landholdings.

⁵³ But one should keep in mind that inequality estimates for Germany and the U.K. from Gómez León and De Jong (2019) should be seen as conservative, as they generally do not account for inequality within a social class. For this reason, total inequality was probably higher in both countries than in prewar Czech Lands. For example, Lindert and Williamson (1983) suggest the U.K. Gini to be 50.2 in 1913.

⁵⁴ Income inequality fluctuated at a relatively high level in Western European countries during the world wars and the interwar period (fluctuations were also more pronounced in the development of top income shares than in more comprehensive measures such as the Gini coefficient). For the U.K., Gómez León and De Jong (2019, Figure 2) suggest that levels of the Gini coefficient during the interwar period (around 40–45) were quite similar to their pre-First World War levels (around 45), and relatively stable. In Germany, after a dramatic decrease during the Weimar period (the Gini fell from 42 in the early 1920s to 30 by 1930), inequality recovered to its prewar levels by the late 1930s (Gini around 42–3). Atkinson (2015) thus broadly indicates that a secular decrease in income inequality in European countries occurred in post-Second World War decades.

⁵⁵ We also drew on pioneering household budget surveys conducted before the Second World War to explore how incomes were utilized and how spending patterns changed over time. We present the results in Appendix D. In summary, data suggest that household consumption patterns stayed surprisingly stable over time. The lion's share of household budgets was spent on basic goods and services, including food, clothing, and housing. Both in capitalism and under communism households consumed most of their income. While the savings rate increased over time, it remained below a sixth of household income (that is, much less than today). Incomes, therefore, continued to play an important role in household decision making under communism.

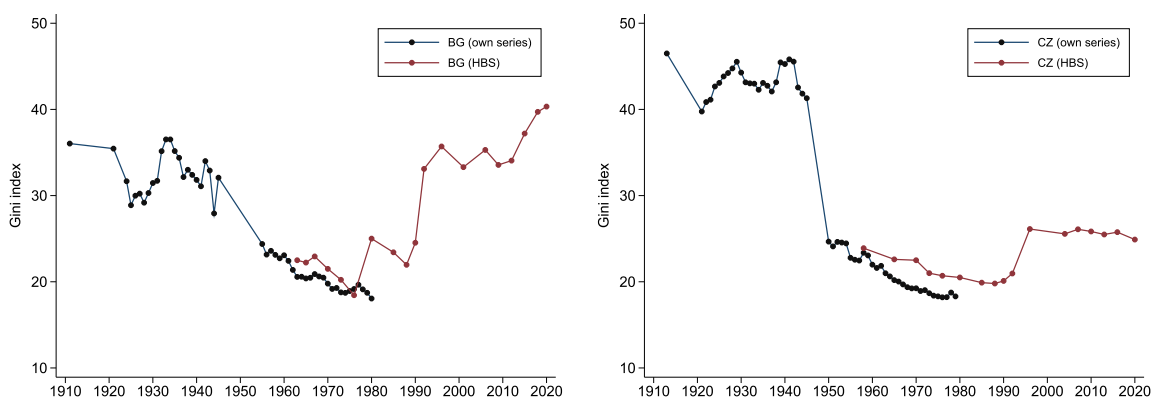


Fig. 13. Long-run evolution of income inequality in Bulgaria and Czechoslovakia/Czech Republic.

Notes: The Gini coefficient for income inequality in Bulgaria (left panel) and in Czechoslovakia/Czech Republic (right panel). Blue lines: our estimates; Red lines: official estimates based on Household Budget Survey (HBS). The HBS series refers to the distribution of the net income per capita. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Source: Own series: see text; HBS series: Milanović, B: All the Ginis (ALG) Dataset at <https://stonecenter.gc.cuny.edu/research/all-the-ginis-alg-dataset-version-february-2019/>; UNU-WIDER, World Income Inequality Database (WIID) Companion dataset (wiidcountry and/or wiidglobal). Version 20 November 2023. <https://doi.org/10.35188/UNU-WIDER/WIIDcomp-281123>.

benchmark estimates to draw more than a century-long picture of the income distribution in Bulgaria and Czechoslovakia/Czech Republic. There are several points worth noting. First, it is reassuring that our estimates based on the social tables approach are remarkably close to HBS estimates in the overlapping years (that is in the 1960s and 1970s). Next, it can be seen that income inequality increased in both countries after the fall of communism in Eastern Europe, suggesting an overall U-shape pattern of inequality development from the beginning of the twentieth century until today. Lastly, and quite interestingly, the post-communist increase in inequality was much more pronounced in Bulgaria than in the Czech Republic (or for that matter in Slovakia). Today, Bulgaria exhibits one of the highest inequality levels in Europe, the Czech Republic one of the lowest. We thus document a striking “reversal of fortune” in inequality dynamics, as Bulgaria used to be much more equal than the Czech Republic, while the opposite is the case today.⁵⁶

7. Conclusion

This paper fills a gap in the literature by providing new long-term estimates on the historical evolution of income inequality in Eastern Europe. We examine inequality during dramatic socio-economic and political changes in the twentieth century, focusing on Bulgaria and Czechoslovakia. These two countries, representing the least and the most developed parts of Eastern Europe at the time, allow us to investigate questions central to inequality and economic development. How does inequality evolve during the process of economic development and which structural forces shape it? How does development impact social structure? And which social groups benefit or lose from economic growth?

Our inequality estimates, based on a newly-constructed dataset and the social tables approach, reveal several important findings. We find that Czechoslovakia had significantly higher inequality than Bulgaria before 1945. Inequality substantially increased during the Great Depression in Bulgaria and fell in Czechoslovakia under German occupation. After the communist revolution, inequality in the two countries quickly converged to similarly low levels. The initial difference in inequality reflected large differences in socio-economic structure. Czechoslovakia had a more industrialized economy and a more stratified society. Our decomposition analysis reveals that inequality within the dominant agricultural sector was key in Bulgaria, whereas inequality both between and within social groups and economic sectors mattered in Czechoslovakia. The swift fall in inequality after the communist ascension to power was driven by collectivization of agriculture, nationalization of industry, and an unprecedented equalization of incomes between social groups.

The experience of Bulgaria and Czechoslovakia suggests that economic forces significantly affected short-run inequality, but institutional and political factors had the largest impact on long-term inequality in the studied Eastern European countries during the twentieth century.

CRediT authorship contribution statement

Stefan Nikolić: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Filip Novokmet:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Piotr Paweł Larysz:** Writing – original draft, Investigation, Data curation.

⁵⁶ Similarly, Russia used to be more equal than Western European countries at the beginning of the twentieth century (Lindert and Nafziger, 2014), while today it is more unequal (Novokmet et al., 2018).

Table A.1
Economic sectors and branches in Bulgaria.

No.	Sector	Branch
1	Primary	Agriculture, horticulture and animal husbandry
2	Secondary	Mining
3	Secondary	Quarrying
4	Secondary	Saltworks
5	Secondary	Metal industry
6	Secondary	Ceramics
7	Secondary	Furniture industry
8	Secondary	Textiles
9	Secondary	Hides, skins and other hard materials derived from animals
10	Secondary	Food and beverages
11	Secondary	Chemicals and related products
12	Secondary	Luxury goods, science, art and letters industry
13	Secondary	Construction industry
14	Secondary	Production and distribution of physical forces
15	Secondary	Construction of transport equipment
16	Secondary	Clothing, footwear and dressing industry
17	Tertiary	Transport and Communication
18	Tertiary	Commerce, credit, insurance
19	Tertiary	Free professions
20	Tertiary	Public administration

Data availability

Data will be made available on request.

Appendix A. Data construction

A.1. Socio-economic structure in Bulgaria

A.1.1. 1910–1946

Based on Bulgarian population census results for 1910, 1920, 1926, 1934, and 1946, we can consistently track the labor force in three social groups: workers, employees, and independents. The independents category include self-employed and their assistants (helping family members). Each social category is further divided according to economic branches and gender. When necessary we aggregate economic branches to ensure comparability across census years. [Table A.1](#) lists the resulting economic branches, spanning all three economic sectors.

Independents and assistants were the most numerous social groups. Together they accounted for around 80% of the labor force (ca. 30% independents and 50% assistants). Most independents were male farmers. Almost all assistants, both men and women, were occupied in agriculture. By and large, they were spouses and children of independent farmers. Other independents were occupied outside of agriculture, for example, as artisans, business owners, or free professionals.

Workers comprised around 15% of the labor force. They were the only social category relatively evenly represented in all three major economic sectors, for both genders. Men were slightly more represented in industry than in other sectors, while women worked marginally more in agriculture. Overall, most workers were men: for each female worker, there were around four male workers.

Employees were the smallest social group. Although their relative size somewhat increased over time, they accounted for at most 5% of the labor force. The tertiary sector occupied practically all employees, both male and female, with few men being employed also in industry. Women were primarily employed in education, whereas men were spread across trade and finance, education, and public administration. As with workers, there were around four times more male than female employees.

To be able to interpolate between census years, we made two additional steps. First, the 1910 and 1920 census report assistants and employees jointly, while subsequent censuses separate these categories. For 1910 and 1920 we disaggregate the “assistants and employees” into two separate categories applying 1926 weights. Second, the 1946 census provides information on the four social groups for the entire economy, but not separately by economic branches. Census data confirm that the overall distribution of the labor force by social groups in 1934 and 1946 was practically the same at the economy level. We therefore keep the branch-level distribution observed in 1934 also for the next census year.

A.1.2. Communist period

The population census of 1956 (Population census, 1956, volume III, pp. 282–318) breaks down the labor force by economic sectors, social classes, and gender. While it is possible to match economic sectors with those reported in previous population censuses, matching social groups is difficult. This is because of both the removal of old and the introduction of new social groups under communism. For this reason, we use the 1956 population census to benchmark total labor force figures by economic sectors, but for

the social categories within these sectors as well as for variation over time we utilize information from statistical yearbooks, which correspond well with pre-communist social categories.

Communism brought major changes in the social structure of the labor force. In 1956, members of labor-cooperative agricultural holdings (TKZS) represented the largest communist social category numbering 1.7 million people and accounting for 42% of the labor force. This was a new social category formed in large part from previously independent farmers and assistants. Accordingly, the share of independent farmers in the labor force shrank to 20% and that of assistants to 13%. Meanwhile, workers became the second largest social category increasing their relative share in the labor force to 23%. Employees were not anymore distinguished from other social categories, but probably bundled together with workers. The remaining 3% of the labor force mainly consisted of artisans in cooperatives and independent artisans. The last two social categories, private traders and free professions, were tiny (0.14% and 0.17% of the labor force).

We combine employment data from the 1956 census and post-war statistical yearbooks in four steps. First, we collect census data on the labor force by economic sectors, social groups, and gender. We match economic sectors with those reported in previous population censuses. Second, we use the totals reported by statistical yearbooks to project backwards and forwards from 1956 the absolute figures in each economic sector. Third, for each year, we distribute the active population in each economic sector into social categories by applying the relative shares for workers, industrial-technical workers, and employees observed in statistical yearbooks. Fourth, we include independent farmers, which are reported in the census. We decrease their number over time according to their share in total agricultural production, reported in statistical yearbooks.

A.2. Incomes in Bulgaria

A.2.1. 1911–1945

Incomes in agriculture. We estimate incomes in agriculture in two steps. We first assign income from labor to each social category. We then estimate income from land and add it to the labor income of independent farmers. Using data on land distribution from agricultural surveys, we differentiate independents and their incomes by land size into eight categories. We reason that independent farmers in the lowest category of land ownership earn little beyond their imputed labor income. Farmers in higher land categories earn more, depending on the size of their plot.

Labor income. Annual incomes of workers in agriculture, from 1924 to 1945, are available from Chakalov (1946, p. 124, table 110). We project annual incomes of agricultural workers backwards, with the annual series of day laborer's wage reported in the Statistical Yearbook (1945, p. 215). We assume that independents and an equivalent number of assistants earned 90% of the income of workers. We consider these assistants to be the spouses of the independents. Remaining assistants, that is other helping family members such as children, earned 70% of the income of workers. To employees, the number of which was negligible, we assign the same income as workers.

Land income. We distribute independents and their assistants to eight social classes and impute land income to each class. We do this in three steps: (1) Distribute the number of independents and their assistants using the distribution of number of farms. We multiply the number of independents and their assistants by the share of farms in each land size class. (2) Distribute income from land using the distribution of farm area. We first subtract our estimated total income from labor from the national income in agriculture reported by Chakalov (1946). We then distribute the income residual to independents and their assistants. To do so, we multiply the total number of independents and their assistants by the share of farm area in each land size class. (3) Calculate income from land per landowner in each land size class. We divide the number of independents and their assistants in each class with the income from land in each class.

Land distribution. Statistical yearbooks report the number of farms and their total area by farm size classes for benchmark years before the First World War, during the interwar, and after the Second World War. Table A.2 reports these data for 1908, 1926, 1934, and 1956 by six farm classes. In 1908, 57% of farms were smaller than 5 hectares. The share of farms below 5 hectares increased to 62% by 1926, and to 68% by 1934. Under communism, 92% of private farms were smaller than 5 hectares. These agricultural small holdings captured respectively 17%, 24%, and 30% of the total privately owned arable land in 1908, 1926, and 1934. By 1956, as much as 81% of private arable land was owned by smallholders. Redistribution of land after the First World War and especially after the Second World War is evident in the statistics. In 1908, largest plots (over 20 hectares) captured 26% of all arable land. In 1926 and 1934, this share fell to 13% and 9%, respectively. By 1956, the few large plots remaining in private ownership held a meager 1% of arable land. We make use of such information on land distribution to account for inequality in agricultural incomes.

Incomes of industrial workers. Our starting point are detailed statistics on daily wages of workers in industry, which come from Statistical Yearbooks. Data are available for 1909, 1911, 1912, 1921, and 1924–1945. Until 1921, an average wage is reported for industry sectors and sub-sectors. From 1924, wages are reported for occupational groups within each industrial sector, separately for male and female workers. Over time, wage statistics expand to include more occupational groups as well as differentiation by age, qualification, or wage type (time or piece wage). By aggregating several sectors, we can consistently track wages across 10 industrial sectors. We take the reported average wage per industrial sector, and average across occupational groups where necessary. We average male wages, which are available for all occupational groups. Female wages, unfortunately, are reported only for certain occupations. We estimate workers' annual incomes by multiplying reported daily wages by 240 days of work, which Chakalov (1946, table 60) uses to annualize daily wages of workers in small-scale industry.

Incomes of other social groups in industry and services. For all social groups in industry and services, except the already discussed industrial workers, we rely on data provided by Chakalov (1946) and Ivanov (2012). We estimate incomes of assistants,

Table A.2
Land distribution in Bulgaria.

Hectares	No. of farms (%)				Arable land (%)			
	1908	1926	1934	1956	1908	1926	1934	1956
0 to 3	41.2	41.2	45.2	69.2	7.5	9.3	11.9	45
3 to 5	15.5	20.1	22.3	23.6	9.4	14.3	18.1	36
5 to 10	24.4	25.7	23.7	7	26.8	34.5	36.8	17.8
10 to 15	10.3	8.1	6	0.1	19	19.8	17.1	1.1
15 to 20	4.2	2.6	1.7	0	10.9	9.5	7.3	0
over 20	4.4	1.8	1.1	0	26.4	12.6	8.8	0

Sources: *Statistical yearbook 1913–1922*, B10, table 7. *Statistical yearbook 1941*, p. 203, table I. *Statistical yearbook 1956 (1957)*, p. 45, table 5.

employees, and independents in both industry and services, by multiplying the wages of workers in a given branch by the relevant skill premia.

In the secondary sector, the skill premium of assistants relates incomes of journeymen with workers in arts and crafts. The skill premium of employees, on the other hand, relates incomes of employees and personnel with workers in big industry. Lastly, the skill premium of employers relates incomes of employers with workers in small-scale industry. These skill premia vary over time, but are constant across industrial branches.

In the tertiary sector, skill premia vary across four branches: transport and communications; commerce, credit, and insurance; liberal professions; and public administration. In commerce, credit, and insurance, we assign the same income to workers and assistants. The available data do not allow separating the incomes of workers and assistants. We calculate the income of employees as the average income of white collars (office personnel; managers and directors) engaged in commerce. To employers we assign an average merchant's income. We calculate the merchant premium as the ratio of employer and worker income in commerce, credit, and insurance.

For social groups in transport and communications, we combine information on income of relevant social categories in automobile and other transport services. We estimate the income of employers by multiplying the income of workers with the merchant premium.

We differentiate incomes of workers and employees in public administration. For workers we apply incomes of municipal servants, while for employees we use income of better-paid, state servants. We do not include assistants and independents in public administration since their number was negligible and because we do not have information on their incomes.

For liberal professions it is not possible to make a distinction of incomes by social classes. Both [Chakalov \(1946\)](#) and [Ivanov \(2012\)](#) only provide the average income for liberal professions. We therefore apply the available average income in each social class.

Gender pay-gap. In both industry and services, we estimate income differences across gender by applying an industry average gender premium. We calculate the gender premium from workers' wages for males and females reported in Statistical Yearbooks. In agriculture, we do not differentiate incomes by gender since we do not have any information on the gender pay gap.

A.2.2. Post-1945

Incomes of workers and employees. Statistical yearbooks provide detailed information on average wages and wage distributions in the state sector. These include state enterprises and cooperatives, but exclude cooperated farmers (TKZS). For mining and industry, we apply the reported branch-level wages of workers (distinguishing industrial-technical workers), and employees. These wage series start in 1955. We match the reported wages with 10 industrial branches in our social table, including mining. To be consistent with our industry classification, we average wages in metallurgy and metal processing as well as in paper and printing, weighting by the labor force shares of these industries.

For workers in state agriculture, construction, and the service sector we combine branch-level average wages and wage distributions to differentiate wages across five income classes. We do this to account for income variation within branches for which we only observe average wages. Starting from the average wage in each branch, we estimate the ratio to the mean for five income classes: bottom 30%, 30%–50%, 50%–70%, 70%–90%, and top 10%. To be precise, we use average wages and corresponding wage distributions for agriculture, forestry, construction, transport, communications, trade, science, education, health, finance, and public administration. To fit our industry classification, we then calculate labor force weighted wages for transport and communication; trade and finance; and science and education. We apply the within inequality estimates for agriculture, forestry, construction, and all five service branches in our social table (see [Table A.1](#)). The average wage is available annually from 1955. We observe wage distributions in 1960, 1962, 1965, 1967, 1971, 1974, 1977, and 1980. We hold the wage distribution constant before 1960 and interpolate for missing years within the observed period.

Incomes of farmers in state cooperatives (TKZS). We first estimate the average income of cooperated farmers by multiplying the average wage of workers in agriculture (reported in Statistical Yearbooks) with the income of cooperated farmers relative to the income of workers reported in household budget surveys.

We then estimate income differences between collectivized farmers, as we did for workers in state agriculture. Starting from the average income, we estimate the ratio to the mean for five income classes: bottom 30%, 30%–50%, 50%–70%, 70%–90%, and top 10%. The average income and wage distributions are available in 1963, 1965, 1967, 1969, 1970, 1972, and 1974. We hold the income distribution constant before 1963 and after 1974, and interpolate for missing years within the observed period.

Incomes of independent farmers. We estimate the income of independent farmers using the same approach as for the pre-communist period. In short, independent farmers earn a fixed labor income and a land income that varies with the size of their farms. We combine information from statistical yearbooks on land distribution and production to classify independent farmers and their land income into five categories. To farmers in each land size category, we impute labor income equal to 90% of a workers wage. We take wages of agricultural workers reported in Statistical Yearbooks. We estimate incomes of independent farmers from 1955 until 1967, after which they do not appear anymore in land statistics. But, because of land collectivization, the number of independent farmers became minuscule already from 1959.

Land income. We distribute independent farmers to five social classes and impute land income to each class. As before, we do this in three steps: (1) Distribute the number of independent farmers using the distribution of number of farms. We multiply the number of independent farmers by the share of farms in each land size class. Farm distributions are available for each year in the period from 1953 to 1956. We take the number of independent farmers reported in the census of 1956. For other years, we estimate the number of independent farmers based on the development of their share in total arable land. (2) Distribute income from land using the distribution of farm area. Statistical yearbooks report the value of agricultural production by social categories. We distribute the total value produced by independent farmers across land size classes. (3) Calculate income from land per landowner in each land size class. We divide the number of independent farmers in each class with the income from land in each class.

Land distribution. Communist collectivization of agriculture, including redistribution of land, was unprecedented. Recall that, in 1934, 68% of farms were smaller than 5 hectares and they accounted for 30% of total privately owned arable land (Table A.2). By 1956, as much as 93% of farms were smaller than 5 hectares and they captured 81% of land owned by independent farmers. This means that the majority of farms owned by independent farmers under communism were small. Moreover, the share of arable land held by independent farmers decreased with collectivization of agriculture. Whereas in 1952 independent farmers owned 40% of arable land, already in 1956 their share fell to 18%. By 1959 independent farmers owned 1% or less of arable land.

A.2.3. Top incomes

Data used to estimate top income shares for Bulgaria come from the statistics of the *Supplementary Tax over Total Income*. These data were reported annually from 1925 to 1946 in the Statistical Yearbook of the Kingdom of Bulgaria. The high exemption level limits our analysis to 0.2% of control population (minimum proportion covered through all years). The statistics comes in the tabulation form, with taxpayers ranged by income brackets according to specific tax rate of the progressive schedule. The tabulations are ranged by the gross income concept (income before personal deductions and income taxes), which is the income concept of our interest. The number of brackets varies over the period from ten to twelve. Each income bracket contains number of tax units and the corresponding tax obligation. The information on total income by bracket was not reported for most of the years. For the period from 1942 to 1945, we have information both on the number of tax units and on the corresponding total income, and, fortunately, Nedkov (1938) reports assessed incomes for 1927 and 1929–1934 (provided by the Ministry of Finance). For years where there is reported only the number of taxpayers, we estimated the corresponding incomes in brackets by assuming Pareto distribution for high incomes.

Top income shares series are constructed by taking the following steps. First, we estimate total income in each bracket of income tax tabulations for years where the number of tax units is reported only and information on the associated income are missing, by assuming that top incomes follow the Pareto distribution. Second, we estimate the control total for population following definitions of tax units specified by tax law and reported in tabulations. The tax unit was defined as household and the total number of households is estimated as the number of adults minus the number of married females. The data are found in population censuses. Third, we estimate the total income denominator. This is based on historical national accounts (Chakalov, 1946; Ivanov, 2012). Finally, we use Pareto interpolation to arrive at income shares of the specific percentages of population.

A.3. Socio-economic structure in Czech Lands/Czechoslovakia

A.3.1. 1910–1950

We consistently track three social groups – workers, employees, and independents – between 1910 and 1950. We collected data from population censuses surveying Czech Lands in 1910, and Czechoslovakia in 1921, 1930, and 1950. We aggregated certain social categories and economic branches to ensure consistency over time. Independents include self-employed and their helping family members. Employees consist of officials and clerks. Workers, laborers, apprentices, and servants comprise the third social group. We distinguish the labor force in these social groups by gender and 23 economic branches listed in Table A.3.

Workers were the largest social group throughout the studied period. They accounted for over half of the labor force until 1920. Their share in the labor force decreased between 1920 and 1950 by 14 percentage points. Most workers were engaged in industry, and the share captured by the secondary sector increased over time. By 1950, three-quarters of male workers and two-thirds of female workers were engaged in industry. The gender balance remained fairly stable over time with at most a third of workers being women.

Between the 1910 and 1921 censuses, there was a notable decline in female employment in agriculture, horticulture, and animal husbandry. The Population census of 1930 recognized prior errors, particularly in the classification of women's occupations. In particular, many wives and daughters of farmers were wrongly classified as “family members without their own occupation” in the 1921 census, despite their active involvement in agriculture, leading to substantial discrepancies in female employment across the years 1910, 1921, and 1930. In light of this, we attributed 32.6% of female dependents in agriculture to female independents, effectively recalibrating the gender balance among independent agricultural workers in 1921.

Table A.3
Economic sectors and branches in Czech lands and Czechoslovakia.

No.	Sector	Branch
1	Primary	Agriculture, horticulture and animal husbandry
2	Secondary	Mining and metallurgy
3	Secondary	Metal processing
4	Secondary	Industrial machinery, tools, and equipment
5	Secondary	Stone and earth industry
6	Secondary	Glass industry
7	Secondary	Chemical industry
8	Secondary	Gasworks, waterworks, power plants
9	Secondary	Wood industry
10	Secondary	Paper industry
11	Secondary	Printing industry
12	Secondary	Textile industry
13	Secondary	Leather industry
14	Secondary	Clothing industry
15	Secondary	Food industry
16	Secondary	Construction
17	Tertiary	Trade
18	Tertiary	Banking
19	Tertiary	Post, telegraph and telephone
20	Tertiary	Railways and other tracks
21	Tertiary	Other transport
22	Tertiary	Judicial and public administration
23	Tertiary	Education and training

The share of employees in the labor force increased significantly over time. Consisting of only 5% of the labor force in 1910, the share of employees started to increase in 1920 and by 1950 reached around 25% of the labor force. The share of officials and clerks increased rapidly after the First World War, with a more pronounced effect on men. This change was especially significant in Trade, Banking and Public Enterprises, likely driven by the introduction of the legal definition of employees. A similar trend was observed in Austria, where the number of employees and civil servants almost tripled from 1910 to 1951, primarily due to the Employees Act of 1921 and civil servant appointments in public companies after 1945 (Mesch and Weigl, 2012). The share of employees in industry increased over time, but the tertiary sector remained the biggest employer of this social group. There were hardly any employees in agriculture. The gender ratio of employees was skewed heavily in favor of men. Although the gender balance somewhat improved over time, even in 1950 there were four male employees for each employed woman.

Independents comprised a stable share of the labor force, between 35% and 40%. Most independents were occupied in agriculture (around two-thirds, even in 1950). The social group was relatively equality split between men and women. However, men were by and large self-employed, whereas women were primarily helping family members. The share of self-employed within the group decreased over time, from around two-thirds before the First World War to one-half in 1950.

A.3.2. Post-1950

The first census taken after the Second World War (1950) in Czechoslovakia is consistent with the pre-war censuses, but not with subsequent ones. Whereas it is possible to match the economic sectors across time, social groups underwent a large change in 1961. For this reason, it is not possible to interpolate population by social groups between 1950 and 1961. Hence, we benchmark our post-1945 estimates on social groups reported in the 1961 census and rely on annual information from statistical yearbooks for variation in the number of people within social groups over time.⁵⁷

The population census of 1961 (Population census, 1961, volume II, pp. 52–69) distinguishes the labor force by economic sectors, social classes, and gender. The labor force counted 6,4 million people. There were seven social classes. Workers (3,3 million) were by far the most numerous. They were followed by 1.8 million “other employees”, 930 thousand farmers in cooperatives, and 95 thousand producers working in non-agricultural cooperatives. Only 1% of the labor force earned a living independent of the state. Among them independent farmers were the most numerous (166 thousand), while self-employed in industry and people with free professions in services were tiny (respectively, 4 and 6 thousand). The low number of independent farmers relative to those in cooperatives suggests that collectivization of agriculture was widespread by 1961.

To be as consistent as possible with pre-war social categories, and in line with the available post-war information on incomes, we define the following social groups after 1950: workers, employees, independents, and members of cooperatives.

To better capture variation within “other employees”, we distinguish officials from engineering-technical employees. Engineers and technicians accounted for a quarter of all employees in the labor force. Their share was especially high in industry, in which they constitute almost a half of all employees. The way we differentiate officials from engineers and technicians is straightforward.

⁵⁷ Federální statistický úrad (1960, 1965, 1967, 1968, 1970, 1971, 1973, 1975, 1977, 1978, 1980, 1982, 1984, 1985, 1986, 1988, 1989).

We take the reported number of engineering-technical employees and apply the branch-specific employee gender ratios to estimate the number of male and female engineers and technicians. We then subtract the number of engineers and technicians from the reported number of employees to calculate the number of officials by gender in each economic branch.

A.4. Incomes in Czech Lands/Czechoslovakia

A.4.1. 1913–1945

Workers (industry and agriculture). Our starting point are detailed statistics on wages of workers in agriculture and industry, which come from Statistical Yearbooks.⁵⁸ In 1913 and 1918–1935, the data cover 15 sectors and 55 sub-sectors. Wages are differentiated by three macro regions: Bohemia; Moravia and Silesia; and Slovakia and Subcarpathian Ruthenia. For the latter, data become available in 1921. We calculate wages for the whole of Czechoslovakia as a population weighted average of regional wages. We estimate annual incomes by multiplying reported daily wages by 300 days of work.

From 1939 to 1945, for the Protectorate of Bohemia and Moravia, we collected hourly wages of workers across 15 industrial sectors from Statistical Yearbooks. We assume an eight-hour workday and then annualize wages as before. For wages in agriculture during the Protectorate, we rely on Stádník (1946).

Employees (industry and agriculture). We estimate the incomes of employees in agriculture and industry by multiplying the wages of workers in a given branch by a skill premium between white-collar and blue-collar workers. The skill premium varies over time, but is constant across sectors.

Our starting point are detailed statistics on salaries of employees in agriculture, industry, commercial trade and banking reported for the period 1929–1936 in Statistical Yearbooks.⁵⁹ For these four sectors, we compute a gender-specific skill premia by dividing the reported salaries of employees by the average income of workers weighted by population. Moreover, we compute skill premium that covers all insured workers and employees, across all sectors, between 1913 to 1945. This cross-sector skill premium serves as a basis for interpolating sector- and gender-specific skill premiums for the years with missing data.

To compute the cross-sector skill premium we rely on several sources. For 1913, we use health insurance data reported by the Imperial and Royal Ministry of the Interior of Austria–Hungary.⁶⁰ It provides daily wages of different types of workers, in all Crown lands, disaggregated at the level of political and judicial districts. For political districts, where wages are disaggregated in court districts and/or different groups of workers, we draw an average for each of the groups defined above. We then assign these wages to the respective number of workers in each group within political districts. We do so in order to calculate aggregated average daily wages of each group in Bohemia, Moravia and Silesia, as well as gender and wage premia.

The skill premia for the period 1929–1943 comes from insurance data reported by Stádník (1946). Stádník reports the total wages and number of manual workers insured at the Central Insurance Company. He also reports the total salaries and number of white-collar workers insured by the General Pension Institute. Based on these data, we calculate the average wage per manual worker and the average salary per white-collar worker. We estimate the skill premium as the ratio between the white- and blue-collar average remuneration.

Before 1929, direct wage estimates for white-collar workers in Czechoslovakia are lacking. To estimate skill premium trends, we link them to wage data for building workers in Germany as reported by Bry (1960). An exception is 1926, where we calculate skill premiums using the average salary for employees from Hotowitz (1927), divided by the weighted average annual wage for all sectors.

Employees (services). For our prewar benchmark, we use Zemská statistická kancelář království Českého (1913), which reports data on the number and salaries of state employees in four social categories, equivalent to those in the social tables.

We use detailed data on state officials and their salaries provided in statistical yearbooks for the years 1924–1947.⁶¹ It includes information on the salaries of state officials and their number, divided by departments of state administration and specific occupations. We were able to match these numbers with our social categories and thereby compute average monthly salaries of independents, officials and clerks, and domestic servants. We approximated the wages of public sector workers based on skill premia estimated on the basis of data reported by Stádník (1946). We complement the missing information on remuneration schemes in education, army and judiciary with data from Landau (1927, 1931). For years where the tax data is missing, we interpolate the wages using the overall wage development reported in statistical yearbooks and other sectors, or employ a linear approximation.

Gender premia. We estimate the incomes of males and females by multiplying the incomes of different social categories by a gender premium. The premium is based on wages of male and female workers. Before the First World War, the premium is calculated from the health insurance data reported by the Imperial and Royal Ministry of the Interior of Austria–Hungary. Statistical yearbooks cover the period from 1928 to 1948.⁶² We apply the same district-level population weighting as we did for skill premia, to calculate the average gender premia for Czech lands in 1913. For the post-war period, we rely on gender differences in workers' pay reported at the regional or national level. As before, when necessary, we weight regional premia by population. For missing years we use linear approximation.

⁵⁸ L'Office de Statistique de la République Tchèque (1928, 1932); Statní úřad statistický (1924, 1932, 1936, 1946, 1948).

⁵⁹ Statistisches Staatsamt (1937, 1938).

⁶⁰ k.k. Ministerium des Innern (1914).

⁶¹ L'Office de Statistique de la République Tchèque (1935); Statistisches Zentralamt (1941, 1942); Statní úřad statistický (1925, 1948); Statistisches Staatsamt (1928, 1934, 1936, 1938).

⁶² Statistisches Staatsamt (1936); Statistisches Zentralamt (1941, 1944); Statní úřad statistický (1948a).

Independent farmers. Land distributions reported for 1902, 1921 and 1930 by Otáhal (1963) and Stadnik's (1946) estimates of farmer's income allow us to differentiate independent farmers and their incomes by land size into six categories. We estimate the income of independent farmers in three steps. First, we multiply the number of independent farmers (this social group also includes tenants and helping family members) by the share of farms in each land size class. This gives us the number of independent farmers in each of the six categories. Second, we distribute Stadnik's (1946) total farmer income to each category based on the share of farm area in each land size class. Third, to the estimated land income in each category we impute Stadnik's (1946) farmer wages to capture labor income.

We thus apply the same approach as we did in Bulgaria. Farmers receive income from land and from labor. The more land they own, the higher their income. Income from labor varies with farmer wages. Farmers with small plots earn little beyond their imputed labor income. By contrast, farmers with large plots earn more from land than from labor.

Independents in industry and services. We follow a different procedure for independents in industry and most services. Based on tax statistics, we are able to differentiate the number of households and their incomes, by the income source and tax brackets, for the secondary (industry and crafts) and tertiary sectors (trade, finance, and transport). Most households in these sectors earned their income from business and self-employment. This income category, accordingly, captures the largest share of total household income. These tax statistics, available for 1928, 1930, 1932, 1933, 1934, 1936, 1938, and 1940, classified households into six income classes. Households that earned below the threshold were excluded.

We estimate the number of independents in industry and services, and impute their income in several steps. First, we estimate the number of tax-paying individuals in each tax bracket, separately for industry and crafts; and trade, finance, and transport. We assume that in both sectors households earning an income from business and/or self-employment had a single adult breadwinner. For this reason, we equate the number of households with the number of tax-paying individuals per income bracket. We then apply gender ratios to differentiate the number of males and females. The gender ratio is based on reported census figures on independents in industry or independents in trade, finance, and transport. We calculate individual income of independents by dividing the total household income from business and self-employment with the estimated number of tax-paying individuals. We subsequently apply the calculated gender premia to estimate incomes of male and female independents in industry and services.

Second, we transfer the estimated number of independents in industry and services and their income to the social tables. We distribute the number of independents observed in the census according to the estimated distribution of tax-paying individuals by income class. Each of the six income brackets forms a sub-category of independents in the social table. We apply the estimated distribution for industry and crafts to all 16 industrial branches in our social table, and the distribution for trade, finance, and transport to four service branches (trade; hostelry; banking; other transport). To account for income variation within industry, we multiply the estimated income of tax-paying individuals in industry and crafts with the ratio of the average wage in a given industrial branch to the average wage across all industrial branches.

Third, we add a seventh sub-category of independents to account for non-fillers, that is individuals that were excluded from tax statistics for earning below a minimum threshold. To be consistent with the census figures, we deduct the total number of estimated, tax-paying individuals from the total number of independents in industry and services reported in the census. This difference gives us the estimated number of non-fillers. We impute the income of non-fillers as 90% of the lowest income category recorded in the tax statistics.

Lastly, we had to estimate incomes of independents in industry and services in the years for which tax statistics are not available. For the gap years in the period between 1928–1940 (1929, 1931, 1935, 1937, and 1939) we used linear interpolation. We used Stadnik's (1946) series on entrepreneurial profits to project incomes forwards from 1940 to 1945. We used Pryor et al.'s (1971) Czechoslovak GDP estimates by sector to project incomes backwards from 1928 to 1921, and to 1913.⁶³ We reflat their constant price GDP with the cost-of-living index from Teichova (1988) and for 1913 with the official wholesale price index reported in postwar statistical yearbooks. As a robustness check, we compared our income total for independents in industry and services with Stadnik's totals for entrepreneurial profits and Nachtigal's (1969) estimated profits in national income in 1913. During the Great Depression (1932–1935), our income totals are around 20% higher than Stadnik's, which is expected since we do not account for unemployment. During the Protectorate, our totals are some 20% lower than Stadnik's, since we do not have data for some economic activities such as banking. To correct for larger discrepancies in 1913 and 1921, we applied an adjustment factor that realigns our estimates with the relevant totals reported in national income studies. Importantly, this adjustment corrects the level of estimated incomes, but not the income distribution of independents in industry and services.⁶⁴ In years not affected by depression or war, our totals align well with Stadnik's.

A.4.2. Post-1945

Workers (industry and agriculture). We collected wages of workers in industry and agriculture from Statistical Yearbooks.⁶⁵ Average wages are available for workers in 15 industrial branches, from 1949. Wages in agriculture start in 1953. We extrapolate agricultural wages from 1950 to 1952 with the overall wage index.

⁶³ Prior to 1929, Stadnik's series on entrepreneurial profits are available only for 1921.

⁶⁴ In particular, for 1913, we applied an adjustment factor of 1.8 that is needed to arrive at a 15% profit share in national income (we assume that the profit share in 1913 was ca. 50% higher than in 1921, which according to Stadnik's figures was 9.5% of national income). For 1921, we applied an adjustment factor of 0.72 to get from our estimated incomes to Stadnik's reported entrepreneurial profit.

⁶⁵ Federální statistický úrad (1960, 1965, 1967, 1968, 1970, 1971, 1973, 1975, 1977, 1978, 1980, 1982, 1984, 1985, 1986, 1988, 1989).

Employees (industry and agriculture). We estimate incomes of employees in industry and agriculture by multiplying the wages of workers in a given branch by a skill premium. We observe the premium of engineering and technical personnel, and of administrative and clerical personnel, relative to manual workers. The skill premium varies over time and is reported separately for construction and all remaining sectors in the Statistical Yearbooks.⁶⁶ We complement several missing years using data from Adam (1984) and Myant (1989), or employ linear approximation.

Independent and cooperated farmers. Statistical yearbooks report the average wage of cooperated farmers and the average wage in agriculture. We estimate the average wage of independent farmers by multiplying the average agricultural wage with the ratio of farmer to worker income reported in budget surveys of agricultural households. We then distribute the average income of independent and of cooperated farmers into five income classes: bottom 30%, 30%–50%, 50%–70%, 70%–90%, and top 10%. We observe wage distributions from household budget surveys taken in 1956, 1958, 1960 1965, 1970, 1973, 1976 and 1980.

Services. For the service sector we combine branch-level average wages and wage distributions to differentiate wages across five income classes. We do this to account for income variation within branches for which we only observe average wages. We utilize branch-specific wage distributions for transport, trade, public administration, healthcare, and education. For banking, we apply the service-sector wage distribution. Starting from the average wage in each branch, we estimate the ratio to the mean for five income classes: bottom 30%, 30%–50%, 50%–70%, 70%–90%, and top 10%. The average wage is available annually from 1948 to 1988 in Statistical Yearbooks.⁶⁷ We observe wage distributions from household budget surveys taken in 1956, 1958, 1960 1965, 1970, 1973, 1976 and 1980.

Gender premia. The 1956 microcensus provides income distributions of workers and of employees, across different sectors, separating males and females (Tables 108–110). The tabulations organize workers and employees in fifteen wage or salary ranges for thirty sectors, providing information on the number of surveyed workers and employees. From these distributions, we calculated mean incomes by gender (we used g-Pinter tool to estimate mean incomes), which we used to calculate gender premia in different sectors (for example, income of males relative to mean income). We then apply these gender premia to average wages in corresponding sectors that we observe from statistical yearbooks. We assume a gradual decline in gender disparities by 1970.

Appendix B. Sources

B.1. Bulgaria

Population censuses.

Direction generale de la statistique. [Glavna direktsiia na statistikata] (1926). Obshti rezultati ot prebroiavane na naselenieto v Tsarstvo Bŭlgariia na 31 dekemvri 1920 god (Résultats généraux du recensement de la population dans le royaume de Bulgarie au 31 décembre 1920). Sofia: Dŭrzhavna pechatnitsa.

Direction generale de la statistique. [Glavna direktsiia na statistikata] (1931). Obshti rezultati ot prebroiavane na naselenieto v Tsarstvo Bŭlgariia na 31 dekemvri 1926 god (Résultats généraux du recensement de la population dans le royaume de Bulgarie au 31 décembre 1926). Sofia: Dŭrzhavna pechatnitsa.

Direction generale de la statistique. [Glavna direktsiia na statistikata] (1937). Prebroiavane na naselenieto na 31 dekemvri 1934: obshti rezultati. Sofia: Dŭrzhavna pechatnitsa.

Tsentralno statisticheskoto upravlenie (1959). Prebroiavane na naselenito na 1 dekemvri 1956 god. : obshti results. Sofia: Nauka i izkustvo

Statistical yearbooks.

Direction generale de la statistique. (1910–1945). Annuaire statistique du Royaume de Bulgarie. Sofia: Dŭrzhavna pechatnitsa.

Direction generale de la statistique. (1947–1980). Annuaire statistique de la Republique populaire de Bulgarie. Sofia: Dŭrzhavna pechatnitsa.

Income studies.

Chakalov, A. (1946). The National Income and Outlay of Bulgaria for 1924–1945, Sofia: Knipegraf.

Ivanov, M. (2012). The Gross Domestic Product of Bulgaria 1870–1945, Sofia: Ciela.

Nedkov, B. (1938). Razvitie na bŭlgarskata finansova sistema prez poslednoto desetilyetie [Development of the financial system of Bulgaria during the last decade], Sofia.

⁶⁶ Federální statistický úrad (1960, 1965, 1968, 1971, 1973, 1975, 1977, 1979, 1980, 1983, 1984, 1986, 1989).

⁶⁷ Federální statistický úrad (1958, 1959, 1960, 1963, 1967, 1969, 1973, 1977, 1979, 1984, 1985, 1989).

B.2. Czechoslovakia

Population censuses.

- Österreichische Statistik (1916). Berufsstatistik nach den Ergebnissen der Volkszählung vom 31. December 1910 in Österreich. Kaiserlich-Königliche Hof- und Staatsdruckerei, Wien.
- Statní úřad statistický (1925). Sčítání lidu v republice Československé ze dne 15. února 1921. Statní úřad statistický : V komisi knihkupectví Bursík a Kohout.
- Statní úřad statistický (1934). Sčítání lidu v republice Československé ze dne 1. prosince 1930. Statní úřad statistický : V komisi knihkupectví Bursík a Kohout.
- Statní úřad statistický (1951). Soutisky obyvatelstva v Československu v letech 1946–1947. Praha, SÚS 1951. 38, 561 s., Čs. statistics, sv. 184 (řada IV., seš. 15).
- Ustřední komise lidové kontroly a statistiky (1965). Sčítání lidu domu a bytu v ceskoslovenské socialistické republice k 1. březnu 1961.

Statistical yearbooks.

- k.k. Ministerium des Innern (1914). Amtliche Nachrichten des k. k. Ministeriums des Innern betreffend die Unfallversicherung und die Krankenversicherung der Arbeiter. Kaiserlich-Königliche Hof- und Staatsdruckerei, Wien.
- L'Office de Statistique de la République Tchèque (1928). Manuel Statistique de la République Tchèque de L'Office de Statistique – Depositaire Bursík & Kohout.
- L'Office de Statistique de la République Tchèque (1932). Manuel Statistique de la République Tchèque de L'Office de Statistique – Depositaire Bursík & Kohout.
- L'Office de Statistique de la République Tchèque (1935). Manuel Statistique de la République Tchèque de L'Office de Statistique. Prague: ORBIS.
- Statistisches Staatsamt (1934–1938). Statistisches Jahrbuch der Tschechoslovakischen Republik. Prag: Obris.
- Statistisches Zentralamt (1941–1944). Statistisches Jahrbuch für das Protektorat Böhmen und Mähren. Prag: Verlag des Statistischen Zentralamtes in Prag, Buchdruckerei Melantrich.
- Statní úřad statistický (1924). Zprávy Statního úřadu statistického republiky Československé. Roč. 5.
- Statní úřad statistický (1932). Zprávy Statního úřadu statistického republiky Československé. Roč. 13.
- Statní úřad statistický (1936). Zprávy Statního úřadu statistického republiky Československé. Roč. 17.
- Statní úřad statistický (1946). Zprávy Statního úřadu statistického republiky Československé. Roč. 27.
- Statní úřad statistický (1948). Statistická ročenka republiky Československé. Prague: Statní úřad statistický.
- Statní úřad statistický (1948a). Statistical Bulletin of Czechoslovakia, 3(9): 159–160.
- Federální statistický úřad (1958–1989). Statistická ročenka Československé socialistické republiky. Prague: Český statistický úřad.
- Federální statistický úřad (1985). Historická statistická ročenka ČSSR. Prague: Český statistický úřad.
- Zemská statistická kancelář království Českého (1913). Statistická příručka království Českého. Prague: Nákladem vlastním - V zakázce knihkupectví Frantiska Řiváče.

Income studies.

- Adam, J. (1984). Employment/Wage Policies in Poland, Czechoslovakia, and Hungary Since 1950. Springer.
- Hotowitz, R. (1927). Naše hospodářství roku 1926 a jeho další vývoj. Česká národohospodářská společnost.
- Landau, Ludwik (1927). Uposażenia urzędników państwowych w Polsce a w Niemczech, Austrii i Czechosłowacji. Statystyka Pracy, zes. 3/4 z. 1927
- Landau, Ludwik (1931). Uposażenia urzędników państwowych w Polsce a w Niemczech, Austrii i Czechosłowacji. Statystyka Pracy, zes. 4 z. 1931
- Maňák, J. (1967). Problematika odměňování české inteligence v letech 1945–1948 (Příspěvek k objasnění počátků nivelizace). Sociologický časopis/Czech Sociological Review, 3(5), 529–540.
- Mesch, Michael (1984). Workers' existence in the late founding period — trade unions and wage development in Austria 1890. Vienna: Europaverlag.
- Myant, M. (1989). The Czechoslovak economy 1948–1988: The battle for economic reform. Cambridge University Press.
- Nachtigal, V. (1969). Národní důchod Československa: dokumentační shrnutí podkladů z období 1913–1966 a analýza období 1946–1955, Praha : Ekonomický ústav ČSAV, 1969
- Pryor, F. L., Pryor, Z. P., Stadnik, M., and Staller, G. J. (1971). Czechoslovak aggregate production in the interwar period. Review of Income and Wealth, 16, 35–59.
- Reif, G. (1931). Vývoj mezd v Čs. republice na léta 1914 až 1930 [Wage Developments in the Czechoslovak Republic for the Years 1914 to 1930]. In Ministerstvo sociální péče (Ed.), Sociální revue, věstník Ministerstva sociální péče. Praha: Ministerstvo sociální péče. Red. Jos. Kotek.
- Stádník, Miloš (1946). Národní Důchod A Jeho Rozdělení, Se Zvláštním Zřetelem K Československu. Prague: Knihovna Sborníku Véd Právních A Statních.

Appendix C. Source reliability

C.1. Bulgaria

Bulgarian historical statistics are generally regarded in the literature to be of high quality and among the best in the region. Detailed economic and population statistics allowed [Chakalov \(1946\)](#) to produce his well-known estimates of Bulgarian national income (1924–1945), and [Ivanov \(2012\)](#) to replicate, improve, and also extend his series back in time (to 1870). Curiously, Bulgaria was among the pioneers of national accounting: it is among the first 10 countries in the world to have developed estimates of national income. Statistics, thus, is one area in which Bulgaria traditionally excelled.

Official Bulgarian population statistics begin to be recorded in 1881, soon after the establishment of the Bulgarian Statistical Office (1880). Five population censuses were carried out already in the period from 1881 to 1905. We rely on the six subsequent population censuses taken in the period 1910–1956. Whereas late nineteenth century censuses have been noted for their deficiencies, best statistical practices were introduced by the turn of the century, following recommendations from the International Statistical Institute. The same methods, ideas and principles laid down by the census in 1900 were kept until the census in 1946 ([Yordanov, 2022](#), p. 129). From 1956, there was a change in the way data were collected. The self-completion principle was abandoned and the interview method was used instead. For us, however, this change has little relevance. More importantly, the 1956 census changed the way data were presented. As we have discussed in detail in [Appendix A](#), social categories in the 1956 census are not fully comparable with those of previous population censuses. As would be expected, new social categories appeared in the census in the communist period.

We dealt with two other comparability issues in the censuses. First, we separated assistants and employees in 1910 and 1920 based on the shares observed in 1926. To be clear, individuals in both social categories were recorded in the same way across the three censuses, but only the latter census reported them as separate categories. There is no reason to believe that the relative number of assistants and employees changed significantly between these censuses. Second, we distributed the 1946 labor force into social groups based on the 1934 census. Both censuses recorded individuals in all four social categories in the same way. The latter census, however, only reported the entire labor force by social categories, without distinguishing social categories by economic sectors. This is why we are forced to rely on the branch-level distribution observed in the previous census. While there may have been some unobserved change in certain economic sectors, the similar distribution of the overall labor force by social categories in the two censuses suggests that there was no economy-wide change.

Statistical yearbooks are our main source of information on incomes in Bulgaria. The first Bulgarian statistical yearbook was published in 1910. Statistical yearbooks started systematically reporting wages of industrial workers already before the First World War. The practice continued after the Great War and carried on largely uninterrupted until 1945. Wages covered workers in private and state-owned establishments. The level of statistical detail increased over time to include a growing number of industrial branches and occupations, distinctions by skill (skilled/unskilled), age (above/under 18), gender, and work type (piece or time). The data are by no means ideal, but they allow us to consistently measure average wages by industrial branches for a relatively long time period.

Statistical yearbooks specify published surveys of industrial enterprises as their source of wage data until 1925. For later years, statistical yearbooks reference internal manuscripts of the Bulgarian Statistical Office as their wage data source. After 1945, statistical yearbooks reported incomes with some gap years from 1948 to 1955, and thereafter continuously for each year. Average yearly salaries are distinguished between many economic sectors and industrial branches, and include both state enterprises and cooperatives. Agriculture is the exception, where state agricultural holdings are covered, but not state cooperatives (TKZS). Unfortunately, after 1945, statistical yearbooks offer no detailed insights on the gender pay gap.

C.2. Czechoslovakia

Czech statistics have a long tradition and are commonly noted for their high quality. In 1856 the Central Committee for Agricultural and Forestry Statistics of Bohemia was established, which is regarded as the commencement of state-organized statistics across Czech lands. Modern population censuses were carried out in a ten-year interval since 1869. Established in 1897, the Statistical Office of the Kingdom of Bohemia was the first centralized statistical body on the territory of the present Czech Republic. Founded in 1919, the State Statistical Office of Czechoslovakia played a crucial role in conducting nationwide statistical surveys, including the population census of independent Czechoslovakia.

Modern population censuses carried out by Austria–Hungary in the late nineteenth and early twentieth centuries were well-organized, providing valuable insights into the population of Central Europe at the turn of the century. The first post-independence census was taken on 15 February 1921. An important difference between the 1910 and 1921 census is the notable decline in female employment in agriculture, horticulture, and animal husbandry. The Population census of 1930 (p. 36) recognized this prior error in the classification of women’s occupations. In particular, many wives and daughters of farmers were wrongly classified as “family members without their own occupation” in the 1921 census, despite their active involvement in agriculture, leading to substantial discrepancies in female employment when compared to 1910 or 1930. In light of this, we attributed the pre-war share (32.6%) of female dependents in agriculture to female independents in 1921, effectively recalibrating the gender balance.

After the Second World War, the new State Statistical Office began surveying the economy amid social changes. [Appendix A](#) provides a detailed discussion of the changes in data reporting by the census during the communist era. We dealt with two

comparability issues. Since 1960, employees were divided into engineering-technical staff and other officials. Similarly, the number of independents was reported separately for cooperatives (farmers, producers) and other independents (farmers, self-employed, free professionals). We estimate gender-specific engineers by applying branch ratios, then calculate officials by subtracting engineers from total employees.

The first Statistical Handbooks of the Kingdom of Bohemia were published in 1909 and 1913, and presented statistical data on Bohemia, comparing it with Moravia, Silesia, and the entire Habsburg monarchy across various domains. After the First World War, the statistical office started systematically reporting average daily wages of insured workers in Statistical Yearbooks. Insured wages were retrieved from the annual Bulletin of the Central Social Insurance Agency (ÚSP). A reform in 1924 (Zákon č. 221/1924 Sb.) extended sickness, invalidity, and old-age insurance to all workers, exempting state officials. This law also initiated the establishment of the ÚSP for managing blue-collar workers' insurance. As pointed out by Stadnik (1946), the wages reported by ÚSP only capture those workers who were employed, thus excluding seasonal workers in agriculture and forestry. The yearbooks were published regularly across the entire studied period with the exception of 1949 to 1956.

As discussed in [Appendix A](#), the incomes of employees and state officials are not reported by the ÚSP. This is because they were insured through a different scheme and obligated to obtain mandatory pension insurance since 1906 through the Employees' Pension Insurance Act. After independence, a reform (Zákon č. 92/1918 Sb.) established the VPÚ for disability and pension insurance of white-collar workers and state employees. A reform in 1920 (Zákon č. 89/1920 Sb.) unified pre-war regulations and a subsequent law from 1929 (Zákon č. 26/1929 Sb.) extended pension insurance to all white-collar employees. To estimate their income before Czechoslovak independence in 1918, we compute skill premia using wage data published by the Imperial and Royal Ministry of the Interior. This data reports the income of workers and employees and was used by [Mesch \(1984\)](#) in his study of pre-war wages.

After the First World War, the statistical office started reporting the average income of employees in statistical yearbooks. Employee incomes were retrieved from the VPU Bulletin. To compute the income of state officials we also use data from statistical yearbooks and reports on legal remuneration schemes from Landau (1927, 1931). For missing years, we compute skill premia wage data for German construction workers from [Bry \(1960\)](#), except for 1926, where skill premiums are calculated using insurance data reported by [Hotowitz \(1927\)](#). Similar to workers, this data covers only those employees who were covered by health insurance.

In the socialist era, statistical yearbooks documented the monthly average earnings of workers in specific branches of the socialist sector of the national economy, derived from register data. These data do not consider personal expenses or transfer payments. To determine the salaries of administrative and clerical personnel, as well as engineering and technical personnel, we utilized information on their overall average earnings within industry. We calculated skill premiums based on these data and subsequently applied calculated premiums to the aforementioned wages of workers.

While industry wages of workers were broadly accessible, those for agriculture, services, and the public sector were reported for the entire workforce, lacking a breakdown into various social categories. We retrieved the data from *Historická Statistická Rocenka* (1985) and complemented it with comparable wages from Statistical Yearbooks (1958, 1959, 1960, 1963, 1967, 1973, 1977, 1979, 1984, 1985, 1989). To differentiate incomes within these sectors, we utilized microcensus data for 1956 and 1970 that classify population within each economic sector by income brackets.

Appendix D. Additional results

D.1. Consumption patterns in capitalism and communism

Household budget surveys (HBS) offer insights into how incomes were utilized and how spending patterns changed over time. We collected information from HBS for several benchmark years that are available in both Bulgaria and Czechoslovakia. Before the Second World War hundreds of households were surveyed, while after the war surveys expanded to include thousands of households. Results of these surveys distinguished households by social groups. We can thus examine expenditure patterns of workers, white collars, and cooperated farmers (who appear in the communist period). [Table D.1](#) presents household budget shares (by four major budget categories) and savings rates, distinguishing between available social groups in Bulgaria and Czechoslovakia for three benchmark periods.

The table suggests three main findings. First, both before and after the Second World War, all social groups spent the vast majority of their budget on food, clothing, and housing. Other expenses accounted for a comparatively smaller share of household budgets. The share of food and clothing reduced over time, but it remained above 50% of the household budget for all social groups. The share of housing fell in Bulgaria after the Second World War, whereas in Czechoslovakia it remained practically unchanged.

Second, household consumption patterns of different social groups converged over time. Before 1945, consumption patterns of blue- and white-collar households somewhat differed. Compared to worker households, white-collar households spent less on food and more on other expenses that included various goods and services (e.g. cultural activities, personal hygiene, transport, but also taxes). Under communism, consumption patterns of different social groups became more similar in Czechoslovakia and practically equalized in Bulgaria.

Third, the lion's share of income, both before and during communism was consumed and not saved. In 1927/28, the savings rate was 4% or less in both Bulgaria and Czechoslovakia for the observed social groups. While the saving rate increased after the Second World War, savings still accounted for less than a sixth of total household income of any observed social group. This is strong empirical evidence that incomes continued to play an important role in the behavior of households under communism.

Table D.1
Household budget shares, by social groups, in Bulgaria and Czechoslovakia.

Bulgaria	1927–28		1963			1975		
	Worker	Assistant	Worker	Assistant	Coop. farmer	Worker	Assistant	Coop. farmer
Food and drink	52%	39%	50%	49%	51%	45%	41%	46%
Clothing	13%	14%	13%	14%	13%	11%	12%	9%
Housing	22%	25%	14%	12%	14%	13%	13%	13%
Other	13%	22%	23%	25%	22%	31%	35%	32%
Savings rate	2%	4%	10%	11%	11%	15%	15%	15%
Czechoslovakia	1927–28		1963			1974		
	Worker	Official	Worker	Employee	Coop. farmer	Worker	Employee	Coop. farmer
Food and drink	55%	39%	42%	37%	34%	30%	27%	27%
Clothing	13%	15%	25%	26%	31%	29%	29%	30%
Housing	12%	15%	11%	13%	9%	11%	13%	10%
Other	20%	31%	22%	24%	26%	30%	31%	33%
Savings rate	4%	4%	7%	8%	14%	10%	11%	16%

Notes: “Housing” includes rent, furnishing, heating, lighting, and related costs. “Other” includes all other expenses for goods and services (e.g. culture, personal hygiene, transport) and financial costs (e.g. taxes, debt payments). Savings rate is savings as percentage of total income.

Sources: Statistical Yearbook of Bulgaria (1931, 1965, and 1977); Zprávy Státního úřadu statistického republiky Československé. (Roč. 12, 1931); Statistická ročenka Československé socialistické republiky (1965 and 1975).

References

- Adam, J., 1984. *Employment/wage Policies in Poland, Czechoslovakia and Hungary Since 1950*. Springer.
- Aitchison, J., Brown, J.A.C., 1957. *The Lognormal Distribution, with Special Reference to Its Uses in Economics*. Cambridge University Press.
- Aldcroft, D.H., 2016. *Europe's Third World: The European Periphery in the Interwar Years*. Routledge.
- Allen, R.C., 2003. *Farm to Factory: A Reinterpretation of the Soviet Industrial Revolution*. Princeton University Press.
- Allen, R.C., 2009. Engels' pause: Technical change, capital accumulation, and inequality in the british industrial revolution. *Explor. Econ. Hist.* 46 (4), 418–435.
- Allen, R.C., 2019. Class structure and inequality during the industrial revolution: lessons from England's social tables, 1688–1867. *Econ. Hist. Rev.* 72 (1), 88–125.
- Alvaredo, F., Atkinson, A., Chancel, L., Piketty, T., Saez, E., Zucman, G., 2016. Distributional national accounts (DINA) guidelines: Concepts and methods used in WID. world. halshs-02794308.
- Anand, S., Kanbur, S.R., 1993. The Kuznets process and the inequality—development relationship. *J. Dev. Econ.* 40 (1), 25–52.
- Atkinson, A.B., 2015. *Inequality: What Can Be Done?*. Harvard University Press.
- Atkinson, A.B., Micklewright, J., 1992. *Economic Transformation in Eastern Europe and the Distribution of Income*. Cambridge University Press.
- Atkinson, A.B., Piketty, T., 2007. *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. Oxford University Press.
- Atkinson, A.B., Piketty, T., 2010. *Top Incomes: A Global Perspective*. Oxford University Press.
- Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the long run of history. *J. Econ. Lit.* 49 (1), 3–71.
- Bell, J.D., 1977. *Peasants in Power: Alexander Stamboliski and the Bulgarian Agrarian National Union, 1899–1923*. Princeton University Press.
- Bolt, J., Van Zanden, J.L., 2020. Maddison style estimates of the evolution of the world economy. A new 2020 update. Maddison-Project Working Paper WP-15.
- Bourguignon, F., Morrisson, C., 2002. Inequality among world citizens: 1820–1992. *Amer. Econ. Rev.* 92 (4), 727–744.
- Braudel, F., 1979. *Civilisation Matérielle Et Capitalisme*, vol. 3, A. Colin.
- Brenner, R., 1982. The agrarian roots of European capitalism. *Past & Present* (97), 16–113.
- Bry, G., 1960. *Wages in Germany*. Princeton University Press.
- Bukowski, P., Novokmet, F., 2021. Between communism and capitalism: long-term inequality in Poland, 1892–2015. *J. Econ. Growth* 26 (2), 187–239.
- Buyst, E., Franaszek, P., 2010. Sectoral developments, 1914–1945. In: Broadberry, S.N., O'Rourke, K.H. (Eds.), *The Cambridge Economic History of Modern Europe: Volume 2, 1870 To the Present*. Cambridge University Press, pp. 208–231.
- Chakalov, A., 1946. *The National Income and Outlay of Bulgaria 1924–1945*. Knipegraf, Sofia.
- Cvrček, T., 2020. *Schooling Under Control: The Origins of Public Education in Imperial Austria 1769–1869*. Mohr Siebeck, Tübingen.
- Dobb, M., 2012. *Soviet Economic Development Since 1917*. Routledge.
- Dumke, R., 1991. Income inequality and industrialization in Germany, 1850–1913: the Kuznets hypothesis re-examined. In: Brenner, Y.S., Kaelble, H., Thomas, M. (Eds.), *Income Distribution in Historical Perspective*. Cambridge University Press, pp. 117–148.
- Erich, A., 1960. *The Soviet Industrialization Debate, 1924–1928*. Harvard University Press.
- Flemming, J.S., Micklewright, J., 2000. Income distribution, economic systems and transition. In: *Handbook of Income Distribution*, vol. 1, Elsevier, pp. 843–918.
- Galor, O., Moav, O., 2004. From physical to human capital accumulation: Inequality and the process of development. *Rev. Econ. Stud.* 71 (4), 1001–1026.
- Gerschenkron, A., 1962. *Economic Backwardness in Historical Perspective, A Book of Essays*. Belknap Press of Harvard University Press, Cambridge.
- Gibrat, R., 1931. *Les Inégalités Économiques*. Sirey.
- Gómez León, M., De Jong, H.J., 2019. Inequality in turbulent times: income distribution in Germany and Britain, 1900–50. *Econ. Hist. Rev.* 72 (3), 1073–1098.
- Gómez León, M., Gabbuti, G., 2024. Wars, depression, and fascism: Income inequality in Italy, 1900–1950. *Documentos de Trabajo EH-Valencia, DT-EHV 2024/01*.
- Grant, O., 2005. *Migration and Inequality in Germany 1870–1913*. Oxford University Press.
- Havelková, B., 2009. Genderová rovnost v období socialismu. In: Bobek, M., Molek, P., Šimíček, V. (Eds.), *Komunistické Právo v Československu: Kapitoly z dějin bezpráví*. Masarykova Univ., Mezinárodní Politologický Ústav.
- İnalçık, H., 1983. *The Emergence of Big Farms, Çiftlik: State, Landlords and Tenants*. Éditions Peeters.
- Itskhoki, O., Moll, B., 2019. Optimal development policies with financial frictions. *Econometrica* 87 (1), 139–173.
- Ivanov, M., 2012. *The Gross Domestic Product of Bulgaria (1870–1945)*. [Brutniyat vateshen produkt na Bulgaria (1870–1945)]. Siela, Sofia.

- Ivanov, M., Ganev, G., 2014. Bulgarian business elite, 1900s–2000s. In: David, T., Westerhuis, G. (Eds.), *The Power of Corporate Networks: A Comparative and Historical Perspective*. Routledge, London, pp. 213–232.
- Ivanov, M., Kopsidis, M., 2023. Industrialisation in a small grain economy during the First Globalisation: Bulgaria c. 1870–1910. *Econ. Hist. Rev.* 76 (1), 169–198.
- Ivanov, M., Tooze, A., 2007. Convergence or decline on Europe's southeastern periphery? Agriculture, population, and GNP in Bulgaria, 1892–1945. *J. Econ. Hist.* 67 (03), 672–703.
- Jenkins, S.P., 1995. Accounting for inequality trends: decomposition analyses for the UK, 1971–86. *Economica* 29–63.
- Kaldor, N., 1957. A model of economic growth. *Econ. J.* 67 (268), 591–624.
- Kalecki, M., 1945. On the Gibrat distribution. *Econometrica* 161–170.
- Kalecki, M., 1993. *Collected Works of Michal Kalecki: Volume IV: Socialism: Economic Growth and Efficiency of Investment*, vol. 4, Oxford University Press.
- Keynes, J., 1920. *The Economic Consequences of the Peace*. Macmillan, London.
- Klein, A., Schulze, M.-S., Vonyó, T., 2017. How peripheral was the periphery? In: O'Rourke, K.H., Williamson, J.G. (Eds.), *The Spread of Modern Industry To the Periphery Since 1871*. Oxford University Press, Oxford.
- Kocka, J., 1981. Capitalism and bureaucracy in German industrialization before 1914. *Econ. Hist. Rev.* 34 (3), 453–468.
- Krejčí, J., 1972. *Social Change and Stratification in Postwar Czechoslovakia*. Columbia University Press.
- Krejčí, J., Machonin, P., 1998. *Czechoslovakia, 1918–92: a Laboratory for Social Change*. Springer.
- Kuznets, S., 1955. Economic growth and income inequality. *American Econ. Rev.* 45 (1), 1–28.
- Kuznets, S., 1966. *Modern Economic Growth: Rate, Structure, and Spread*. Yale University Press New Haven.
- Kuznets, S., 1973. *Modern economic growth: findings and reflections*. *American Econ. Rev.* 63 (3), 247–258.
- Lakner, C., Milanovic, B., 2016. Global income distribution: From the fall of the Berlin Wall to the Great Recession. *World Bank Econ. Rev.* 30 (2), 203–232.
- Lampe, J.R., 1986. *The Bulgarian Economy in the Twentieth Century*. Croom Helm, London.
- Lampe, J., Jackson, M., 1982. *Balkan Economic History, 1550–1950: From Imperial Borderlands to Developing Nations*. Indiana University Press, Bloomington.
- Lewis, W.A., 1954. Economic development with unlimited supplies of labour. *Manch. Sch.* 22 (2), 139–191.
- Lindert, P.H., 2000. Three centuries of inequality in Britain and America. In: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*, vol. 1, Elsevier, pp. 167–216.
- Lindert, P.H., Nafziger, S., 2014. Russian inequality on the eve of revolution. *J. Econ. Hist.* 74 (03), 767–798.
- Lindert, P.H., Williamson, J.G., 1982. Revising England's social tables 1688–1812. *Explor. Econ. Hist.* 19 (4), 385–408.
- Lindert, P.H., Williamson, J.G., 1983. Reinterpreting Britain's social tables, 1688–1913. *Explor. Econ. Hist.* 20 (1), 94–109.
- Lindert, P.H., Williamson, J.G., 2013. American incomes before and after the Revolution. *J. Econ. Hist.* 73 (3), 725–765.
- Lindert, P.H., Williamson, J.G., 2016. *Unequal Gains: American Growth and Inequality Since 1700*. Princeton University Press.
- Lydall, H., 1968. *The Structure of Earnings*. Clarendon Press, Oxford.
- Maňák, J., 1967. Problematika odměňování české inteligence v letech 1945–1948 (příspěvek k objasnění počátků nivelizace). *Sociologický Časopis/Czech Sociol. Rev.* 529–540.
- Marcus, N., Nikolić, S., Straumann, T., 2021. Economic policy, 1918–1939. In: Morys, M. (Ed.), *The Economic History of Central, East and South-East Europe*. Routledge, pp. 188–215.
- Mesch, M., 1984. *Arbeiterexistenz in der Spätgründerzeit: Gewerkschaften und Lohnentwicklung in Österreich, 1890–1914*. (33), Europaverlag.
- Michal, J.M., 1973. Size-distribution of earnings and household incomes in small socialist countries. *Rev. Income Wealth* 19 (4), 407–427.
- Mijatović, B., Zavadžil, M., 2023. Serbia on the path to modern economic growth. *Econ. Hist. Rev.* 76 (1), 199–220.
- Milanovic, B., 2018. Towards an explanation of inequality in premodern societies: the role of colonies, urbanization, and high population density. *Econ. Hist. Rev.* 71 (4), 1029–1047.
- Milanovic, B., 2023. *Visions of Inequality: from the French Revolution to the End of the Cold War*. Harvard University Press.
- Milanovic, B., Lindert, P.H., Williamson, J.G., 2011. Pre-industrial inequality. *Econ. J.* 121 (551), 255–272.
- Milanovic, B., Mijatović, B., 2021. The real urban wage in an agricultural economy without landless farmers: Serbia, 1862–1910. *Econ. Hist. Rev.* 74 (2), 424–448.
- Modalsli, J., 2015. Inequality in the very long run: inferring inequality from data on social groups. *J. Econ. Inequal.* 13, 225–247.
- Morrisson, C., 1984. Income distribution in East European and Western countries. *J. Comp. Econ.* 8 (2), 121–138.
- Nove, A., 1961. *The Soviet Economy: An Introduction*. Allen & Unwin, London.
- Novokmet, F., 2017. *Entre Communisme Et Capitalisme. Essais Sur L'évolution Des Inégalités De Revenus Et De Patrimoines En Europe De L'est, 1890–2015* (Ph.D. thesis). École des hautes études en sciences sociales.
- Novokmet, F., 2023. Between private and public property: Capital transformations and long-run inequality in the Czech Lands, 1898–2015. mimeo.
- Novokmet, F., Piketty, T., Zucman, G., 2018. From Soviets to oligarchs: inequality and property in Russia 1905–2016. *J. Econ. Inequal.* 16, 189–223.
- Nurkse, R., 1953. *Problems of Capital Formation in Underdeveloped Countries*. Blackwell, Oxford.
- Olšovský, R., 1961. *Přehled Hospodářského Vývoje Československa V letech 1918–1945*. Státní nakladatelství politické literatury, Praha.
- Otáhal, M., 1963. *Zápas O Pozemkovou Reformu V ČSR*. Nakl. Československé akademie věd.
- Palairot, M., 1997. *The Balkan Economies C. 1800–1914: Evolution Without Development*. Cambridge University Press.
- Piketty, T., 2001. *Les Hauts Revenus En France Au Xxe Siècle*. Hachette.
- Piketty, T., 2003. Income inequality in France, 1901–1998. *J. Polit. Econ.* 111 (5), 1004–1042.
- Piketty, T., 2006. The Kuznets curve: Yesterday and tomorrow. In: Banerjee, A.V., Benabou, R., Mookherjee, D. (Eds.), *Understanding Poverty*. Oxford University Press, Oxford, pp. 63–72.
- Piketty, T., 2020. *Capital and Ideology*. Harvard University Press.
- Prucha, V., 2004. *Hospodářské a Sociální Dějiny Československa 1918–1992: 1 Díl. Období 1918–1945*. Doplněk, Brno.
- Ranis, G., Fei, J.C., 1961. A theory of economic development. *American Econ. Rev.* 533–565.
- Ravallion, M., 2016. *The Economics of Poverty: History, Measurement, and Policy*. Oxford University Press.
- Rodríguez Weber, J., 2015. *Income Inequality in Chile Since 1850* (Ph.D. thesis). Programa de Historia Económica, FCS, Udelar.
- Roine, J., Waldenström, D., 2015. Long-run trends in the distribution of income and wealth. In: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*, vol. 2, Elsevier, pp. 469–592.
- Rosenstein-Rodan, P.N., 1943. Problems of industrialisation of eastern and south-eastern Europe. *Econ. J.* 53 (210–211), 202–211.
- Scitovsky, T., 1966. An international comparison of the trend of professional earnings. *American Econ. Rev.* 56 (1/2), 25–42.
- Státník, M., 1946. *Národní Důchod a Jeho Rozdělení: Se Zvláštním Zřetelem K Československu*. Prague: Nákl. Knihovny Sborníku Věd Právních a Státních.
- Teichova, A., 1988. *The Czechoslovak Economy 1918–1980*. Routledge, London.
- Tilly, R.H., Kopsidis, M., 2020. *From Old Regime to Industrial State: a History of German Industrialization from the Eighteenth Century to World War I*. University of Chicago Press.
- Van Zanden, J.L., Baten, J., Foldvari, P., Van Leeuwen, B., 2014. The changing shape of global inequality 1820–2000; Exploring a new dataset. *Rev. Income Wealth* 60 (2), 279–297.
- Večerník, J., 1991. Earnings distribution in Czechoslovakia: Intertemporal changes and international comparison. *Europ. Sociol. Rev.* 7 (3), 237–252.
- Vinski, I., 1967. The distribution of Yugoslavia's national income by social classes in 1938. *Rev. Income Wealth* 13 (3), 259–284.
- Wiśniewski, J., 1934. *Rozkład Dochodów Według Wysokości W R. 1929*. Instytut Badania Konjunktury Gospodarczych i Cen, Badania nad dochodem społecznym w Polsce. t. III. Warszawa.
- Yordanov, A., 2022. Historical and legal analysis of the census of the population in Bulgaria in the period 1879–2010. *Law J. New Bulg. Univ.* 18 (2), 121–136.